

North American Climate in CMIP5 Experiments.

Part II: Evaluation of 20th Century Intra-Seasonal to Decadal Variability

Justin Sheffield, Suzana J. Camargo, Rong Fu, Qi Hu, Xianan Jiang, Nathaniel Johnson,
Kristopher B. Karnauskas, Jim Kinter, Sanjiv Kumar, Baird Langenbrunner, Eric
Maloney, Annarita Mariotti, Joyce E. Meyerson, J. David Neelin, Zaitao Pan, Alfredo
Ruiz-Barradas, Richard Seager, Yolande L. Serra, De-Zheng Sun, Chunzai Wang, Shang-
Ping Xie, Jin-Yi Yu, Tao Zhang, Ming Zhao

Justin Sheffield, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ
Baird Langenbrunner, Joyce E. Meyerson, J. David Neelin, Department of Atmospheric and Oceanic
Sciences, University of California Los Angeles, CA
Suzana J. Camargo, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY
Rong Fu, Jackson School of Geosciences, University of Texas at Austin, TX
Qu Hu, School of Natural Resources and Department of Earth and Atmospheric Sciences, University of
Nebraska-Lincoln, Lincoln, NE
Xianan Jiang, Joint Institute for Regional Earth System Science and Engineering, University of California,
Los Angeles, CA
Kristopher B. Karnauskas, Woods Hole Oceanographic Institution, Woods Hole, MA
Sanjiv Kumar, Jim Kinter, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite
302, Calverton, MD
Eric D. Maloney, Department of Atmospheric Science, Colorado State University, Fort Collins, CO
Annarita Mariotti, National Oceanic and Atmospheric Administration, Office of Oceanic and Atmospheric
Research (NOAA/OAR), Silver Spring, MD
Zaitao Pan, Saint Louis University, St. Louis, MO
Alfredo Ruiz-Barradas, Department of Atmospheric and Oceanic Science, University of Maryland, College
Park, MD
Richard Seager, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY
Yolande L. Serra, Department of Atmospheric Science, University of Arizona, Tucson, AZ
De-Zheng Sun, Cooperative Institute for Environmental Studies/University of Colorado & NOAA/Earth
System Research Laboratory, Boulder, CO
Chunzai Wang, Physical Oceanography Division, NOAA Atlantic Oceanographic and Meteorological
Laboratory, Miami, FL
Jin-Yi Yu, Department of earth System Science, University of California, Irvine, CA
Nathaniel Johnson, International Pacific Research Center, SOEST, University of Hawaii at Manoa,
Honolulu, HI
Shang-Ping Xie, International Pacific Research Center, and Department of Meteorology. SOEST,
University of Hawaii at Manoa, Honolulu, HI
Tao Zhang, Cooperative Institute for Environmental Studies/University of Colorado & NOAA/Earth
System Research Laboratory, Boulder, CO
Ming Zhao, NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ

43

Submitted July 30, 2012

44 *Corresponding author address: Justin Sheffield, Department of Civil and Environmental
45 Engineering, Princeton University, Princeton, NJ, 08540. Email: justin@princeton.edu

46

46 **Abstract**

47 This is the second part of a three-part paper on North American climate in CMIP5 that
48 evaluates the 20th simulations of intra-seasonal to multi-decadal variability and
49 teleconnections with North American climate. Overall, the multi-model ensemble does
50 reasonably well at reproducing observed variability in several aspects, but does less well
51 at capturing observed teleconnections, with implications for future projections examined
52 in part three of this paper. In terms of intra-seasonal variability, almost half of the models
53 examined can reproduce observed variability in the eastern Pacific and most models
54 capture the midsummer drought over Central America. The multi-model mean replicates
55 the density of tropical disturbances and storms but with large spread among the models.
56 The coarse resolution of the models means that tropical cyclone frequencies are under
57 predicted in the Atlantic and eastern North Pacific. The frequency and mean amplitude of
58 ENSO are generally well reproduced, although teleconnections with North American
59 climate are widely varying among models and only a few models can reproduce the east
60 and central Pacific types of ENSO and connections with US winter temperatures. The
61 models capture the spatial pattern of PDO variability and its influence on continental
62 temperature and West coast precipitation, but less well for the wintertime precipitation.
63 The spatial representation of the AMO is reasonable but the magnitude of SST anomalies
64 and teleconnections are poorly reproduced. Multi-decadal trends such as the warming
65 hole over the central-southeast US and precipitation increases are not replicated by the
66 models, indicating that observed changes are linked to natural variability.

67

1. Introduction

This is the second part of a three-part paper on the Climate Model Intercomparison Project phase 5 (CMIP5; Taylor et al., 2012) model simulations for North America. This second part evaluates the CMIP5 models in their ability to replicate the observed variability of North American continental and regional climate, and related climate processes. The first part (Sheffield et al., 2012) evaluates the representation of the climatology of continental and regional climate features. The third part (Maloney et al., 2012) describes the projected changes for the 21st century.

The CMIP5 provides an unprecedented collection of climate model output data for the assessment of future climate projections as well as evaluations of climate models for contemporary climate, the attribution of observed climate change and improved understanding of climate processes and feedbacks. As such, these data will feed into the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5), and other global, regional and national assessments.

The goal of this study is to provide a broad evaluation of CMIP5 models in their depiction of North American climate and associated processes. It synthesizes and draws from individual work by investigators within the CMIP5 Task Force of the US National Oceanic and Atmospheric Administration (NOAA) Modeling Analysis and Prediction Program (MAPP). This is part of a Journal of Climate special collection on North America in CMIP5 and we draw from individual papers within the special issue, which provide more detailed analysis that can be presented in this synthesis paper.

We begin in Section 2 by describing the CMIP5, providing an overview of the models analyzed, the historical simulations and the general methodology for evaluating

the models. Details of the main observational datasets to which the climate models are compared are also given in this section. The next 5 sections focus on different aspects of North American climate variability, organized by the time scale of the climate feature. Section 3 covers intraseasonal variability with focus on variability in the eastern Pacific Ocean and summer drought over the southern US and Central America. Atlantic and east Pacific tropical cyclone activity is evaluated in Section 4. Interannual climate variability is assessed in Section 5. Decadal variability and multi-decadal trends are assessed in Sections 6 and 7, respectively. Finally, the results are synthesized in Section 8.

2. CMIP5 Models and Simulations

2.1. CMIP5 Models

We use data from multiple model simulations of the “historical” scenario from the CMIP5 database. The CMIP5 experiments were carried out by 20 modeling groups representing more than 50 climate models with the aim of further understanding past and future climate change in key areas of uncertainty (Taylor et al., 2012). In particular, experiments have been focused on understanding model differences in clouds and carbon feedbacks, quantifying decadal climate predictability and why models give different answers when driven by the same forcings. The CMIP5 builds on the previous phase (CMIP3) experiments in several ways. Firstly a greater number of modeling centers and models have participated. Secondly, the models are more comprehensive in terms of the processes that they represent and are run at higher spatial resolution, therefore hopefully resulting in better skill in representing current climate conditions and reducing uncertainty in future projections. Table 1 provides an overview of the models used. The specific models used vary for each individual analysis because of data availability at the

time of this study, and so the model names are provided within the results section where appropriate.

2.2. Overview of Methods

Data from the “historical” CMIP5 scenario are evaluated, which is a coupled atmosphere-ocean mode simulation that is forced by historical estimates of changes in atmospheric composition from natural and anthropogenic sources, volcanoes, greenhouse gases and aerosols, as well as changes in solar output and land cover. For certain climate features we also analyze model simulations from the CMIP3 that provided the underlying climate model data to the fourth assessment report (AR4) of the IPCC. Several models have contributed to both the CMIP3 and CMIP5 experiments, either for the same version of the model, or for a newer version, and this allows a direct evaluation of changes in skill in individual models as well as the model ensemble.

Historical scenario simulations were carried out for the period from the start of the industrial revolution to near present: 1850-2005. Our evaluations are generally carried out for the past 30 years, depending on the type of analysis and the availability of observations. For some analyses the only, or best available, data are from satellite remote sensing which restricts the analysis to the satellite period, which is generally from 1979 onwards. For other analyses, multiple observational datasets are used to capture the uncertainty in the observations. The observational datasets are summarized in Table 2 and further details of the datasets and data processing are given in the relevant subsections and figure captions. Where the comparisons go beyond 2005 (e.g. 1979-2008), data from the model RCP8.5 future projection scenario simulation are appended to the model historical time series. About half the models have multiple ensemble members and

these are averaged where appropriate or used to assess the variability across ensemble members.

3. Tropical Intraseasonal Variability

3.1. MJO-related variability over the eastern Pacific and adjoining regions

It has been well documented that convection over the eastern Pacific (EPAC) ITCZ and neighboring areas is characterized by pronounced intraseasonal variability (ISV) during boreal summer (e.g., Knutson and Weickmann, 1987; Kayano and Kousky, 1999; Maloney and Hartmann, 2000a; Maloney and Esbensen, 2003, 2007; de Szoeke and Bretherton, 2005; Jiang and Waliser, 2008, 2009, 2011). ISV over the EPAC exerts broad impacts on regional weather and climate phenomena described in the first part of this paper (Sheffield et al., 2012), including tropical cyclone activity over the EPAC and the Gulf of Mexico, the summertime gap wind near the Gulfs of Tehuantepec and Papagayo, the Caribbean Low-Level Jet and precipitation, the mid-summer drought over Central America and Mexico (see below), and the North American monsoon (e.g., Magana et al., 1999; Maloney and Hartmann, 2000b; Maloney and Hartmann, 2000a; Maloney and Esbensen, 2003; Lorenz and Hartmann, 2006; Serra et al., 2010; Martin and Schumacher, 2010).

Here, model fidelity in representing ISV over the EPAC and Intra-America Sea (IAS) region is assessed by analyzing daily output of rainfall and 850hPa winds from sixteen CMIP5 models. While model-simulated seasonal mean rainfall patterns on continental scales in these models have been examined in Section 3.1 of part I of this paper, detailed regional summer mean rainfall fields over the EPAC are further evaluated

to explore the role of mean state on model performances in simulating the ISV. Figure 1 displays a Taylor diagram for summer mean (May-September) precipitation from in CMIP5 models over the EPAC domain (150°W-80°W; 5°S-30°N) compared to the TMPA precipitation. While the two HadGEM models (HadGEM2_CC and HadGEM2_ES) display the highest pattern correlations (~ 0.93), the MRI_CGCM3 show the smallest RMS due to its better skill in simulating the spatial standard deviations of summer mean rainfall over the EPAC. In addition, four models (MPI_ESM_LR, CSIRO_MK3, CanESM2, and CNRM_CM5) exhibit relatively better pattern correlation scores than other models.

The leading ISV modes over the EPAC based on observed and simulated rainfall fields are identified using a complex empirical orthogonal function (CEOF) approach (Maloney et al., 2008). CEOF analyses are applied to 30-90-day band-pass filtered daily rainfall anomalies and the spatial amplitude and phase for the first CEOF mode (CEOF1) based on TMPA are illustrated in Figures 2a and 2b. Similar to Maloney et al. (2008), the maximum amplitude of the observed rainfall CEOF1 occurs over the far eastern part of the EPAC. Figure 2b illustrates the pattern of spatial phase of observed rainfall CEOF1. Note that the direction of propagation associated with the CEOF1 is indicated by the gradient of the spatial phase. In agreement with previous studies, the observed leading ISV mode associated with the CEOF1 largely exhibits an eastward propagation, while a northward component is also evident (e.g., Jiang and Waliser, 2008; Maloney et al., 2008; Jiang et al., 2011).

Next, the fidelity of the CMIP5 models in simulating the leading EPAC ISV mode is assessed by calculating pattern correlations of the simulated rainfall CEOF1

against observations. To increase sampling, spatial patterns of rainfall anomalies associated with the CEOF1 based on both observations and model simulations are derived at two quadratic phases by multiplying the CEOF1 amplitude by the Cosine and Sine of spatial phase at each grid point, respectively. The pattern correlations are then calculated at both of these two quadratic phases. A final pattern correlation score for a particular model is derived by averaging these two pattern correlation coefficients. Figure 2c illustrates pattern correlation scores in depicting the CEOF1 rainfall pattern for each model simulation versus domain averaged CEOF1 amplitude relative to observations. A majority of the CMIP5 models tend to underestimate the amplitude of the leading EPAC ISV mode associated with the rainfall CEOF1, except CNRM_CM5, MPI_ESM_LR, and HadGEM2_CC and HadGEM2_ES. Seven models exhibit relative higher pattern correlation scores (> 0.75).

The models with relative better skill in representing the leading EPAC ISV mode also largely exhibit better skill for summer mean rainfall (cf. Fig. 1 and Fig. 2c) and 850hPa wind patterns (not shown). A common feature among the more skillful models is the presence of westerly or very weak easterly mean winds over the EPAC warm pool region, as in the observations. Most of the models with relatively lower skill exhibit a stronger easterly summer mean flow (> 4 m/s). This suggests that realistic representation of the mean state could be conducive for improved simulations of the EPAC ISV, which has also been discussed for MJO simulations over the western Pacific and Indian Ocean (e.g., Kim et al., 2009). One hypothesis is that a realistic mean state produces the correct sign of surface flux anomalies relative to intraseasonal precipitation, which helps to destabilize the local intraseasonal disturbance (e.g. Maloney and Esbensen, 2005).

3.2. Mid-summer Drought over Central America

The rainy season in Central America and southern Mexico spans roughly May through October. For most of the region, the precipitation climatology features maxima in June and September and a period of reduced rainfall during July-August known as the midsummer drought (MSD; Portig et al., 1961, Magaña et al., 1999). The MSD is regular enough to be known colloquially and plays an important role in farming practices (Osgood et al., 2009). A previous assessment of CMIP3 model performance at simulating the MSD and future projections (Rauscher et al., 2008) suggested that many models are capable of simulating the MSD despite an overall dry bias, and that the MSD is projected to become stronger with an earlier onset. In this section, the CMIP5 performance at simulating summertime precipitation and the MSD is evaluated. We evaluate 23 CMIP5 models against the TMPA, GPCP and UNAM observational datasets. A simple algorithm for detecting and quantifying the climatological MSD is used that does not assume *a priori* which months are maxima and which months constitute the MSD (Karnauskas et al., 2012).

The large-scale distribution of summertime precipitation across the Inter-Americas region (Figure 3, top panel) is reasonably well simulated by the CMIP5 multi-model ensemble (MME) mean. However, several areas of disagreement exist including underestimates of precipitation over the southernmost Pacific coast of Mexico through Guatemala, the Caribbean coast of Central America, the Yucatan peninsula, Florida, as well as the major Caribbean islands of Cuba, Hispaniola, and Jamaica. Many of these biases may be related to orographic influences such as island effects and gap winds.

Nevertheless, the CMIP5 MME does reasonably capture the essence of the MSD over much of the Inter-Americas region (Figure 3, bottom panel). The maximum strength of the MSD in the MME is found just offshore of El Salvador and represents a midsummer precipitation minimum that is ~ 2.5 mm/day less than the early- and late-summer peaks. Significant differences in the location and strength of the MSD between various observational data sets preclude a definitive evaluation of the CMIP5 MME, but it is clear that the strength of the MSD is underestimated in some regions, including along the Pacific coast of Central America, the western Caribbean, the major Caribbean islands and Florida.

4. East Pacific and Atlantic tropical storm track and cyclone activity

Tropical storms and cyclones can have large social and economical impacts around the globe. It is well known since the 1970s that climate models are able to simulate tropical cyclone-like storms (e.g. Manabe et al., 1970; Bengtsson et al., 1982). As the resolution of the climate models increases, the modeled storm characteristics become more realistic (e.g. Zhao et al., 2009). High-resolution climate models projections for global tropical cyclone frequency and intensity are very robust across models, while still differing on regional scales (Knutson et al., 2010). Analysis of CMIP3 model storms showed that the tropical cyclone-like storms produced still had many biases common of low-resolution models (Walsh et al., 2010). Therefore, various dynamical and statistical techniques for downscaling tropical cyclone activity using only the CMIP3 large-scale variables were employed (Emanuel et al., 2008; Knutson et al., 2008).

Here we analyze the storms in CMIP5 models using two approaches. First, the Hodges (1995) method identifies storm tracks based on 6-hourly 850hPa positive vorticity centers with a minimum threshold of $0.5 \times 10^{-6} \text{ s}^{-1}$ which persist for at least 2 days and have tracks of at least 1000 km in length. This method primarily identifies westward moving disturbances such as easterly waves (e.g. Serra et al., 2010), although more intense storms which could potentially reach hurricane intensity are not excluded. The second approach is based on the tracking method described in Camargo and Zebiak (2002), which uses low-level vorticity, surface winds, surface pressure, and atmospheric temperature to identify tropical cyclones, and considers only warm core storms. Only a subset of the tropical disturbances will intensify enough to be identified by the second tracking method and the percentage that this occurs will vary among different models. The CMIP5 standard models have trouble simulating the number of tropical cyclones, which can be attributed in part to their coarse resolution. In section 4.3 we show results from the GFDL high-resolution model.

4.1. Tropical Storm Track

The tropical storm track density is calculated based on the method of Hodges (1995) using 6-hourly, 850 hPa relative vorticity from the first ensemble members of seven models and compared to the ERA-Interim (Figure 4, top). Mean track strength is simply the mean of the smoothed 850 hPa vorticity along the track (Figure 4, bottom). The multi-model mean track density is in good agreement with the ERA-Interim data, however significant differences are seen with the individual models. The most apparent discrepancies are with the CanESM2 and CCSM4 models, which strongly overestimate

activity across the East Pacific and western Caribbean and suggest a more longitudinally oriented track shifted south from what is observed. HadGEM2-ES and MIROC5 underestimate tracks in the West Atlantic, while GFDL-ESM2M underestimates tracks throughout the region except near 130°W. MPI-ESM-LR also underestimates tracks across the region as well as shifts their location southward. The track density maximum off the west coast of Mexico is best captured by HadGEM2-ES. The mean track strength maximum in the East Pacific lies along the west coast of Mexico and is well captured by the multi-model mean, although it is somewhat stronger than in the ERA-Interim (Figure 4, bottom). On the other hand, the multi-model mean strength in the Gulf of Mexico and West Atlantic along the east coast of the US is strongly underestimated compared to ERA-Interim. Unlike for track density, these biases are fairly consistent among the models.

4.2. Tropical Cyclones in the North Atlantic and Eastern North Pacific

Tropical-cyclone (TC) type structures were tracked in five models for 1950-2005 using the method of Camargo and Zebiak (2002), which uses 6-hourly low-level vorticity, surface winds, surface pressure, atmospheric temperature. The method considers only warm core storms and uses model-dependent (and resolution) thresholds and storms have to last at least two days. We compare with the observed tropical cyclones from best-track datasets of the National Hurricane Center. Figure 5 shows the tracks of tropical cyclone-like structures for five models and the observations for the north Atlantic and eastern north Pacific. The number of TCs in all models is much lower than in observations, which is common to many low-resolution global climate models

(e.g. Camargo et al., 2005, 2007). The HadGEM2 has the largest low bias and the MPI model has the most realistic tracks in the Atlantic basin. The MRI model tracks in the Atlantic are mostly in the subtropical region, with very few storms in the deep tropics. In contrast, in the eastern North Pacific the MRI has storm activity too near the equator. In the eastern North Pacific, very few storms (in all models) have westward tracks. The models seem to have an easier time in producing storms that are in the northwestward direction parallel to the Central America coast.

Figure 6 shows the mean number of TCs per month for the North Atlantic and eastern North Pacific. In some cases, the models produce too many storms in the off-season, while all models produce too few storms in the peak season. The bottom panels show the spread of the number of storms per year as box plots in the models and observations, emphasizing the low number of storms per year in all models. Examination of variability across ensemble members in producing tropical cyclones was carried out for five member runs of the MRI model (not shown) but was much less than among different models.

4.3. High-Resolution Model Results

The ability of models to simulate tropical cyclone activity is limited partly by their relatively coarse spatial resolutions. Recent studies suggest that when forced by observed SSTs and sea-ice concentration, a global atmospheric model with a resolution ranging from 50km to 20km can simulate many aspects of TC/hurricane frequency variability for the past few decades during which reliable observations are available (e.g., Sugi et al. 2002; McDonald et al. 2005; Yoshimura et al. 2006; Oouchi et al. 2006;

Bengtsson et al. 2007; Gualdi et al. 2008; LaRow et al. 2008; Zhao et al. 2009). This includes the geographical distribution of storm genesis and track locations, seasonal cycle and interannual variability of storm frequency over the North Atlantic and East Pacific basins. The success is not only a direct evaluation of model capability but also an indication of the dominant role of SST variability on TC/hurricane frequency variability. In this section we present results from high-resolution simulations with the GFDL C180HIRAM model forced by observed boundary conditions. Historic downscaling experiments (Held et al. 2012) using the model were forced by prescribed inter-annually and seasonally varying SSTs and sea-ice concentration from the HadISST dataset. The 3-member ensemble simulation was run for 1981 to 2008. The storm detection and tracking algorithm is described in (Zhao et al. 2009). We focus on TCs with near-surface wind speed reaching hurricane intensity.

Figure 7 shows the observed and C180HIRAM simulated hurricane tracks for the N. Atlantic and E. Pacific for 1981-2008. The annual mean numbers of hurricanes and their statistics are shown in Figure 8a. The simulations reproduce fairly well the observed statistics although the model appears to slightly underestimate the observed variances of interannual variability in the N. Atlantic. Figure 8b,c show that the model also captures the observed seasonal cycle in both the N. Atlantic and E. Pacific. The model can also reproduce the observed year to year variation of annual hurricane count as well as the decadal trend for both basins for this period (Zhao et al. 2009; Held and coauthors 2012). The quality of the model's present-day simulation of N. Atlantic and E. Pacific hurricane statistics increases our confidence in future predictions using this model.

5. Interannual to decadal variability

5.1. *El Niño-Southern Oscillation (ENSO)*

The ENSO is the most important driver of global climate variability on inter-annual time scales. It impacts many regions worldwide through climate teleconnections (Ropelewski and Halpert, 1987), which link the tropical Pacific to higher latitudes through shifts in mid-latitude weather patterns. The impact of ENSO on North American climate is felt most in the wintertime, with El Niño events bringing warmer temperatures to much of the northern part of the continent and wetter conditions in the southern US and northern Mexico. La Niña events tend to bring drier weather to the southern US. The ability of CMIP5 models to simulate ENSO is carried out for several aspects of ENSO variability and for teleconnections with North American climate.

A. Evaluation of ENSO teleconnections

We examine how well the historical simulations of CMIP5 models reproduce the composite near-surface air temperature (SAT) and precipitation patterns over North America during El Niño and La Niña episodes. In both model and observed data, we define ENSO episodes similarly to the Climate Prediction Center (CPC). A monthly ENSO index is calculated from detrended and high-pass filtered SSTs over the Niño3.4 region (5°S – 5°N, 170°W - 120°W) from ERSSTv3b observations and CMIP5 models. An El Niño (La Niña) episode is defined as any sequence of months where the three-month running mean Niño3.4 SST, is greater than 0.5°C (less than -0.5°C) for at least five consecutive three-month running seasons and the peak amplitude of the ENSO episode

occurs during the boreal fall or winter. In the ensemble mean, the frequency of ENSO episodes and the mean peak amplitude are similar to observed values (not shown).

In observations approximately 93% of El Niño and 90% of La Niña episodes feature peak amplitudes in fall or winter. In the CMIP5 ensemble of the historical simulations, however, only 69% of El Niño and 66% of La Niña episodes have peak amplitudes in fall or winter, although a couple of models (CanESM2 and NorESM1-M) do have fall/winter peak frequencies exceeding 80% for both El Niño and La Niña episodes. This finding suggests that CMIP5 models do not fully reproduce the phase-locking of ENSO to the seasonal cycle, a deficiency noted in CMIP3 models as well (Guilyardi et al. 2009). The following analysis focuses on those episodes that do peak in fall or winter.

For each identified ENSO episode, we calculate composites of seasonal North American SAT and precipitation anomalies (seasonal cycle removed and detrended) from the models and CRU TS3.1 observations, beginning in the fall of the developing episode year, and ending in the summer of the decay year. To focus on differences in spatial patterns rather than differences owing to variations in ENSO amplitude, the anomalies are normalized by the peak amplitude of the episode so that the composite maps represent anomalies per °C of Niño3.4 SST peak amplitude. We assess how well the models reproduce the observed ENSO-related SAT and precipitation patterns if they are statistically distinguishable in a test of field significance based on the “false discovery rate” (FDR) (Benjamini and Hochberg 1995, Wilks 2006) at the 90%, 95%, and 99% confidence levels (Figure 9). Model performance in reproducing the observed patterns is quite variable, with significant differences for many models and seasons and for both

variables and ENSO phases. The ensemble mean calculations (pooled ENSO episodes from all models) reveal that the most significant differences occur in winter and spring. Moreover, because the individual model patterns are quite variable, the amplitude of the ensemble mean pattern is quite weak (see Fig. 10). This point is illustrated in the observed and CMIP5 ensemble composite SAT and precipitation patterns for DJF(0/1), shown in Fig. 10. The observed SAT and precipitation composites feature peak amplitudes exceeding $1^{\circ}\text{C }^{\circ}\text{C}^{-1}$ and $0.5 \text{ mm day}^{-1} ^{\circ}\text{C}^{-1}$, but the CMIP5 ensemble composites are much weaker. Moreover, the composites for La Niña episodes fail even to resemble the general observed patterns. These results suggest that large model differences in ENSO SST patterns, downstream teleconnection patterns, and associated influences on temperature and precipitation result in widely diverging SAT and precipitation patterns associated with ENSO, especially in winter and spring.

B. East Pacific/Central Pacific ENSO and Teleconnections with US Winter Surface Air Temperature

It has been increasingly recognized that different types of ENSO occur in the tropical Pacific (e.g. Wang and Weisberg, 2000; Trenberth and Stepaniak, 2001; Larkin and Harrison, 2005; Yu and Kao, 2007; Ashok et al., 2007; Kao and Yu, 2009; Kug et al., 2009). Two particular types that have been emphasized are the Eastern-Pacific (EP) type that produces SST anomalies near the South America coast and the Central-Pacific (CP) type that produces anomalies near the international dateline. While the EP ENSO is the conventional type of ENSO, the CP ENSO has gradually increased its occurrence during the past few decades (e.g. Lee and McPhaden, 2010). Recent observational studies have

indicated that the impacts produced by these two types of ENSO on North American climate can be different (e.g., Mo 2011, Yu et al. 2012). Here the ENSO teleconnection over the US simulated in the CMIP5 models are further examined according to the ENSO type. Following Kao and Yu (2009) and Yu and Kim (2010), a regression-Empirical Orthogonal Function (EOF) analysis is used to identify the CP and EP types of El Niño from monthly SSTs. The SST anomalies regressed with the Niño1+2 SST index were removed before the EOF analysis was applied to obtain the spatial pattern of the CP ENSO. Similarly, we subtracted the SST anomalies regressed with the Niño4 SST index before the EOF analysis was applied to identify the leading structure of the EP ENSO. The principal components of the leading EOF modes obtained from this analysis represent the ENSO strengths and are defined as the CP ENSO index and the EP ENSO index. The winter (DJF) SAT anomalies regressed to these two indices are different over the US (Figure 11a,b) with a warm northeast to cold southwest pattern for the EP El Niño and a warm northwest to cold southeast pattern for the CP El Niño. Adding these two impact patterns together results in a pattern that resembles the classical warm-north, cold-south pattern of El Niño impact as shown in Figure 10. Data from the NCEP-NCAR reanalysis and ERSST V3b during 1950-2010 were used for this analysis.

We repeated the EOF and regression analyses to evaluate how well the CMIP5 models reproduce the different US impacts to the two types of ENSO. The regressed winter SAT anomaly patterns calculated from twenty CMIP5 models are shown in Figure 11. The observed impact patterns on the US winter SAT are well simulated by some models, such as the MIROC5 and MRI-CGCM3 for the EP ENSO and the NorESM1-M and HadGCM2-ES for the CP ENSO. However, some models show an impact pattern

that is almost opposite to that observed, such as HadCM3 for the CP ENSO and INMCM4 for the EP ENSO. To quantify how well the impact patterns are simulated, pattern correlation coefficients are calculated between the model regressed patterns and the NCEP regressed patterns. As shown in Figure 12, there are seven CMIP5 models (GFDL-ESM-2M, HADGEM2-ES, IPSL-CM5-MR, MIROC5, MPI-ESM-LR, MPI-ESM-P, NorESM1-M) whose pattern correlation coefficients for both the EP ENSO and the CP ENSO are larger than 0.5. This group of the CMIP5 models is considered as the models whose regressed US winter temperature patterns are close to the observed patterns for the two types of ENSO. According to Kim and Yu (2012), five of these models (GFDL-ESM2M, HadGEM2-ES, MPI-ESM-LR, MPI-ESM-P, and NorESM1-M) are also the CMIP5 models that produce strong intensities of both types of ENSO.

C. ENSO warm/cold events asymmetry

ENSO asymmetry refers to the fact that the two phases of ENSO are not mirror images of each other (Burgers and Stephenson, 1999). The asymmetry shows up in both the surface and subsurface fields (Rodgers et al., 2004; Schopf and Burgman, 2006; Sun and Zhang, 2006; Zhang et al., 2009). Causes for such an asymmetry are not yet clearly understood, but accumulating evidence suggests that it is likely a consequence of nonlinearity of ocean dynamics (Jin et al., 2003; Sun 2010, Liang et al., 2012). Asymmetry is also linked to the time-mean effect of ENSO (Sun and Zhang, 2006; Schopf and Burgman, 2006; Sun, 2010, Liang et al. 2012). Understanding the causes and consequences of ENSO asymmetry may hold the key to understanding decadal variability in the tropics and beyond (Rodgers et al., 2004; Sun and Yu, 2009, Liang et al., 2012).

Figure 13 shows the sum of the SST anomalies between the warm and cold phases of ENSO from HadISST observations and CMIP5 models. The threshold value used for defining the warm and cold phase anomalies is set as $+0.5^{\circ}\text{C}$ and -0.5°C respectively. This sum has also been called the SST anomaly residual and has been a common measure of the ENSO asymmetry in the SST field. All models underestimate the observed positive SST residual (and therefore the asymmetry) over the eastern Pacific. Measured by the skewness of Niño3 SST anomalies (which is a more rigorous measure of asymmetry), all the models also underestimate the observed ENSO asymmetry (Figure 14). The figure also shows that the stronger variability of ENSO (measured by variance) does not guarantee a stronger asymmetry in ENSO (measured by skewness).

5.2. Persistent droughts and wet spells over Great Plains and the southern-tier states

Persistent dry and wet summers are features of the US Great Plains and southern US. We evaluate how the CMIP5 models describe the processes that cause such persistent anomalies in terms of low-level circulation and moisture flux anomalies by comparing with the NCEP-NCAR reanalysis. This complements the evaluations of the average seasonal circulation in the region, such as the low-level southerly jet as shown in part 1 of this paper (Sheffield et al., 2012). Persistent wet and dry summers are defined by JJA precipitation anomalies averaged over the Great Plains region from 90° - 105°W and 30° - 50°N during 1971-2000. Wet (dry) summers are identified as having normalized JJA precipitation larger (smaller) than 0.6 (-0.6) standard deviation. The reanalysis data identify 8 wet and 7 dry summers in 1971-2000, and the models identify between 7 and 12 events. We show the composites of vertically integrated moisture from the surface to

top of the troposphere, the 850hPa geopotential height, and near surface winds at 925hPa for the wet and dry summers and their differences for the reanalysis (Figure 15) and for a single model, CCSM4, as an example (Figure 16).

Comparisons of the two figures indicate some similarities and also some quite different processes causing the persistent wet or dry summers. The integrated moisture fluxes in both datasets indicate a high moisture in an averaged cyclonic rotation in the troposphere in persistent wet summers (Figs. 15a and 16a) but anticyclonic rotation in dry summers (Figs. 15b and 16b) in the Great Plains. However, the sources of the moisture and the low-level dynamic structure are quite different. For the reanalysis, the convergence of moisture in the central Great Plains during wet summers results from southerly flow anomalies in the enhanced subtropical high pressure system in the North Atlantic and northerly flow anomalies in low pressure anomalies centered in the Midwest (Fig. 15d). These anomalies suggest a frontal system along the depression from the Midwest to the Southwest. A nearly reversed pattern of flow anomalies is shown during the dry summers (Figs. 15e and 15f). The model simulations show a different pattern of flow anomalies (Figs. 16d and e). In wet summers, the moisture is primarily from the east along the easterly and southeasterly quadrants of a high pressure anomaly center in the Great Lakes areas, instead of from the south as in the reanalysis result (Figs. 16a vs. 15a). In dry summers, the model suggests dry flows from the Mexican plateau off the Sierra Madre Oriental in Mexico. These contrasts are shown in Fig. 16f. The other models in CMIP5 also simulated different tropospheric circulation patterns from those in the reanalysis for either the wet or dry summers in the Great Plains.

6. Decadal Variability

6.1. Pacific Decadal Oscillation (PDO) and its influence on North American climate

On interdecadal timescales, variability in the tropical and extratropical North Pacific, particularly that of the Pacific Decadal Oscillation (PDO), has significant physical and ecological impacts over North America (Mantua et al. 1997, Higgins et al. 2000, Meehl et al. 2012). We examine the PDO and its relationships with N. American temperature and precipitation for 17 CMIP5 models. In both observations and the historical runs we define the PDO in a similar manner to that of Zhang et al. (1997) and Mantua et al. (1997). We define the PDO as the leading empirical orthogonal function of extended winter (November-April) monthly mean sea surface temperature (SST) anomalies in the North Pacific poleward of 20°N (Zhang et al. (1997) and Mantua et al. (1997)) for the period from 1900-1993. We subtract the monthly global mean SST prior to the analysis. We then calculate the PDO index by projecting monthly North Pacific SST anomalies onto the PDO pattern for all available months and then standardizing the resulting time series. For the observed PDO, we use the HadISST dataset for the period between 1870 and 2009. For the CMIP5 models, the analysis period begins as early as 1850 and extends to 2005. We use a single realization for each model. Figure 17 illustrates the PDO patterns in both observations and the CMIP5 ensemble obtained by regressing the unfiltered monthly SST anomalies upon the PDO index for all calendar months. As in the CMIP3 models (Oshima and Tanimoto 2009, Furtado et al. 2011), the CMIP5 models reproduce the basic PDO horseshoe SST pattern. The most notable difference is the westward shift of the North Pacific center of action in models with

respect to observations (Figure 17c). The regions with the largest differences also correspond with regions of relatively high inter-model variability (Figure 17d).

PDO relationships with North American climate are evaluated in Figure 18, which summarizes which CMIP5 regression patterns are significantly different from those based on observations. For each set of regressions, we determine whether the model and observation regression maps are significantly different with a “false discovery rate” (FDR) field significance test, as described in the ENSO composite analysis above. For most models and most seasons, the differences in the regression patterns are not statistically significant. The lack of significant differences likely owes to a combination of small effective sample size, a relatively modest impact of the PDO on seasonal SAT and precipitation, and the ability of the models to capture the general PDO behavior. When considering the entire ensemble, however, the increased sample size brings out differences in winter and spring precipitation regressions that are statistically significant at the 95% confidence level.

To illustrate the discrepancies during DJF, Figure 19 shows the SAT and precipitation regressions in observations and the CMIP5 ensemble. The CMIP5 models do rather well in capturing the PDO influence on North American SAT, with positive (negative) SAT anomalies in northwest (southeast) North America during the positive phase of the PDO. Almost all local differences in the regression coefficients are not statistically significant. In contrast, the CMIP5 models perform somewhat poorly in reproducing the precipitation patterns over large parts of North America. Both observations (Fig. 19b) and CMIP5 ensemble (Fig. 19d) produce a tripole pattern of precipitation anomalies over the west coast of North America. Large differences,

however, are found in eastern North America. In observations, the positive phase of the PDO is associated with reduced wintertime precipitation in the Tennessee and Ohio Valleys, northeastern U.S., and Southeastern Canada (Fig. 19b), but the CMIP5 ensemble fails to discern this influence (Fig. 19d, f). Though of smaller magnitude, significant differences also occur in central North America (Fig. 19f). In spring (MAM) the largest differences in the precipitation regressions occur along the coast of British Columbia, where observed regressions indicate positive anomalies but the CMIP5 ensemble produces a pronounced negative anomaly (not shown). Both observations and the CMIP5 ensemble reproduce positive precipitation anomalies along the west coast and central plains of the U.S.

6.2. Atlantic Multidecadal Oscillation (AMO)

The Atlantic Multidecadal Oscillation (AMO) is an important feature of North Atlantic variability with a period of the order of 65 - 70-yr (Kerr, 2000), and is associated with impacts on tropical and mid-latitude climate, including N. American summertime precipitation (Enfield et al, 2001; Sutton and Hodson, 2005) and linkages to Atlantic hurricane occurrence (Goldenberg et al., 2001). The AMO index is defined as the detrended North Atlantic SST during the Atlantic hurricane season of June to November (JJASON) from the equator to 60°N, 75°W-5°W with the 11-year running mean (e.g., Enfield et al. 2001; Knight et al. 2005). As shown in Figure 20a, the individual models show highly varying amplitudes and phases, but all models capture the North Atlantic warming in the recent decades. Two major discrepancies between the observation and models are during the 1930s-40s when the models underestimate the warming, and

during 1910s-20s when the models underestimate the cooling. The lagged autocorrelations of the AMO indices show the persistence of the modeled AMO (Fig. 20b), suggesting the potential for predicting the North Atlantic SST anomalies.

The tropospheric vertical wind shear in the main development region (MDR) is an important factor that affects the formation and development of Atlantic hurricanes, with a small (large) wind shear favoring (disfavoring) Atlantic hurricane activity. Following a typical wind shear definition in the literature (e.g., Goldenberg et al., 2001; Wang et al., 2006; Aiyyer and Thorncroft, 2006; Vecchi and Soden, 2007), the vertical wind shear is defined as the magnitude of the vector difference between winds at 200 hPa and 850 hPa. The AMO-related wind shear is shown by regressing the vertical wind shear onto the AMO indices during JJASON (Fig. 21). The 20CR reanalysis product shows that the AMO is associated with negative wind shear regressions in the MDR, indicating that the warm (cold) phase of the AMO corresponds to reduced (enhanced) wind shear which is in turn favorable (unfavorable) for Atlantic hurricane activity (e.g., Goldenberg et al., 2001). Of the 23 models, 10 models are able to simulate the negative wind shear in the MDR. The discrepancies in the other models may be due to the inaccurate representation of the AMO itself and/or inaccurate representation of the wind response to the AMO.

The decadal modulation of SSTs in the Atlantic Ocean, via the AMO, has been shown to have an important impact on extreme droughts and wet episodes over the central US (Nigam et al. 2011). The ability of the models to capture the AMO and its hydroclimate impact over North America in summer and fall are explored Figure 22 and shown in more detail in Kavvada et al. (2012). Its SST signature and hydroclimate impact in summer and fall are obtained by regressing it on the regional seasonal precipitation and

SSTs for 1901-1999. The SST signature of the AMO is stronger in fall than in summer and this is reflected in its impact on central US precipitation in observations. In both seasons the SST anomalies reach a maximum over the mid Atlantic, over the sub-polar gyre region, with normal conditions over the Gulf of Mexico and north of the Greater Antilles. The warm phase of the AMO induces drying conditions over the central US and wet conditions over Florida and the US northeast in both seasons, but with more intensity in fall. However, there are seasonally contrasting conditions along the Gulf of Mexico states where decreased precipitation occurs in summer but increased precipitation occurs in fall. These hydroclimate impacts highlight the importance of the AMO seasonality.

In general the models do not capture the SST seasonality of the AMO well. Simulated SST anomalies are larger in summer than in fall and the simulated hydroclimate impact lacks the drying over the central US. The extension of the SST anomalies in CCSM4 is confined to the mid and northern latitudes of the Atlantic. The GFDL-CM3 SST anomalies extend to the whole Atlantic and a footprint over the Greenland Sea is also evident. The SST anomalies in HadGEM2-ES are larger than observations and are placed to the southeast of Greenland. Its structure is similar to that in observations but with a larger area of normal conditions over the western tropical Atlantic in summer. The MPI-ESM-LR model misplaces the maximum SST anomalies over the eastern mid Atlantic, with no observed tropical extension. For the hydroclimate impacts, the buildup of the drying conditions over the central US and the wet conditions along the coastal south Atlantic US states from summer to fall are not well captured by the models. The initial drying over the south central US in summer is only captured by HadGEM2-ES, while the wet conditions over the south Atlantic US states are captured to

different degree by the models. The intensification of the drying of the central US in fall is also best captured by the HadGEM2-ES model, followed by the MPI and GFDL-CM3 models. CCSM4 shows no drying over the region but it is the only model showing wet conditions along the eastern coastal plains of the US.

7. Multidecadal Trends

7.1. Trends in Temperature and the ‘warming hole’ over the southeastern US

A unique of feature of US temperature change during the 20th century is the so-called “warming hole (WH)” observed in the southeastern US (Pan et al, 2004). While global warming accelerated in the 20th century, the WH region experienced cooling, especially in summer during the latter half of the century. Studies have attributed the mechanisms for this abnormal cooling (lack of warming) trend to large-scale decadal oscillations such as PDO and AMO (Robinson et al., 2002, Kunkel et al., 2006, Meehl et al., 2012) and to regional scale hydrological processes (Pan et al., 2004) and land surface interactions (Liang et al. 2007). It was a challenge for most CMIP3 models to reproduce the WH (Kunkel et. al., 2006).

We evaluate whether the CMIP5 models capture the warming hole in Figures 23 and 24, which shows the annual and seasonal trends, respectively, in near surface air temperature from the observation and the CMIP5 multi-model mean. We evaluated 19 models, and for all available ensemble members, totaling 96 ensemble members. For the multi-model mean spatial plot, the best ensemble member from each model having highest spatial correlation with the observation are selected. Model and observation data

are re-gridded to a common resolution $2.5^{\circ} \times 2.5^{\circ}$ using area averaging. Trends are calculated for the 1930-2004 period using the Theil-Sen approach (Theil, 1950; Sen, 1968). The choice of 1930-2004 gives a prominent warming hole signal starting from the warmest decade following the Dust Bowl drought. The models do well in simulating the north-south gradient in warming rates. The multi-model mean also shows a lesser warming region, compared to the surrounding region, in the central US. However, the multi-model mean fails to capture the absolute negative temperature trend as found in the observation.

The observed WH is more evident in the eastern US during summer and thus we separated temperature trend between eastern and western US during summer and winter. (Figure 24). The observed summer cooling is prominent in the north central US, while winter cooling is prominent along the Southeastern Coastal region. In summer, the models generally have difficulty in simulating the cooling trend in the north central US, and also the east-west gradient in summer warming rates (Fig. 24a). In winter, the multimodel mean produces the spatial pattern of the temperature trend, showing less warming in the southeastern coastal region. The east-west gradient in temperature trend also matches reasonably well with the observation (Fig. 24b). The seasonal difference in the models' skill may reflect different mechanisms responsible for the WH in summer and winter. In winter, decadal variation is stronger over the US, whereas in summer local forcing, such as land surface processes, are more prominent. This seems to suggest that models can more easily capture large-scale dynamics in winter including decadal variations, but struggle to simulate regional scale processes during the summer.

Figure 25 shows the temporal evolution of the winter trends over the eastern US in the observation and models, and relative to the western US. The absolute negative trend (warming hole) in the eastern US shows multidecadal variability, which peaked between 1930 and 1960. For most of the 20th century the western US warmed faster than the east (negative East - West Trend). The 95% model spread range brackets the observed multidecadal variability in the eastern US temperature trend. Approximately half of the model simulations show negative temperature trends in the eastern US, as well as a negative (East – West) temperature trend difference. However, the decadal variability varies considerably between models, making it difficult for the multimodel mean to capture the amplitude of observed multidecadal variability. This result suggests that the US warming hole is partly contributed by the natural multidecadal variability in the climate system.

7.2. Trends in Diurnal Temperature Range (DTR)

Observed warming during the day and night has been asymmetric, with nocturnal minimum surface air temperature (Tmin) rising about twice as fast than daytime maximum temperature (Tmax) during the second half of 20th century, mostly during 1950-1980 (Vose et al., 2005). Changes in cloud cover, among atmospheric water vapor, soil moisture and other factors, accounts for 25-50% of the DTR reduction (Dai et al., 1999). Cloud cover, soil moisture, precipitation, and atmospheric/oceanic teleconnections account for up to 80% of regional variance over 1901-2002. Over the U.S., cloud cover alone accounts for up to 63% of regional annual DTR variability (Lauritsen and Rogers, 2012). Global Tmin increased by 0.20 °C dec⁻¹ while Tmax increased by 0.14 °C dec⁻¹

from 1950–2004, resulting in a DTR decrease of -0.07°C (Voss et al., 2005). During the same period over North America ($175\text{--}60^{\circ}\text{W}$, $15\text{--}75^{\circ}\text{N}$), summer T_{max} and T_{min} increased 0.07 and 0.12°C , respectively, resulting in a -0.05°C decrease in DTR. A similar decrease (-0.06°C) occurred in winter. Over the WH region ($105\text{--}80^{\circ}\text{W}$, $30\text{--}45^{\circ}\text{N}$), summer T_{max} actually decreased sharply (-0.13°C) while T_{min} increased slightly (0.05°C), yielding a DTR decrease of 0.18°C . Winter DTR also decreased by 0.13°C .

Figure 26 shows a comparison of the linear trend in DTR from 23 models against the CRU TS3.1 observational dataset. In summer (top), the models capture a cooling trend in central U.S., roughly corresponding to the WH region (Fig. 24) although the position is shifted to the west and the magnitude is lower. In winter the models largely reproduce the broad decrease in DTR as observed, but again with lower magnitude. The inter-model variance (contours in Fig. 26), a measure of uncertainty in the simulations, is largest over southwestern/Mexico (summer) and western (winter) U.S., likely arising from complex topography along the Rocky mountain ranges.

7.3. Trends in Precipitation

Precipitation has generally increased over North America in the last half of the 20th century (Karl and Knight, 1998; Zhang et al., 2000). Trends in precipitation are positively correlated with streamflow trends, thereby affecting water resource availability and flood potential (Lettenmaier et al. 1994; McCabe and Wolock 2002, Kumar et al. 2009). Figure 27 shows the multi-model ensemble average precipitation trend from 19 models against the CRU observations. The multi-model average weakly captures the wetting trend in North America, particularly at higher latitudes. However, the multi-

model average fails to capture the trend magnitude, for example, the higher wetting trend (>20 mm/decade) in the eastern US. Figure 28 (a) and (b) show the 30 year running trend during the 20th century in the eastern and western US, respectively. The 95% model spread brackets the observed precipitation trend magnitude in both regions. The higher wetting trend in the observations has slowed down in the last decade in the eastern US. The muted magnitude of the trend in Figure 27 seems to be a result of low signal to noise ratio (the multi-model median line hovers around the zero line in Figure 28), rather than a robust feature of CMIP5 climate models. Some individual models capture very well the observed trend magnitude (Figure 28). Drying in Mexico is a dominant but incorrect feature in the CMIP5 simulations; a similar feature was also found in CMIP3 climate simulation (IPCC, 2007).

Changes in seasonal precipitation were also calculated for winter and summer but as the difference between the second-half (1951–99) and the first-half- (1901–50) of the 20th century (Figure 29). A similar analysis for CMIP3 models of spring and summer climatological precipitation showed limited success by the models in simulating the observed changes (Ruiz-Barradas and Nigam, 2010). Observations show that winter precipitation over the Pacific Northwest, the coastal states of the southeastern US, and the northeast of the Great Lakes over Canada has increased, but in areas like southern California and the inner states of the southeastern US and the Baja Peninsula precipitation has decreased. In summer, precipitation has decreased over the Southeastern and Northeastern US, but has increased for regions around the Great Lakes, and northeast into Canada, and over parts of Mexico. Those changes are barely captured for any of the CMIP5 models. Among the main changes captured by the models are the increase in

precipitation over the Pacific Northwest in winter, which is captured by the CCSM4 and MPI-ESM-LR models, and the drying over Florida and eastern US in summer which is modestly captured by CCSM4 and GFDL-CM3.

8. Discussion and Conclusions

This paper has evaluated the simulated variability from the CMIP5 multi-model ensemble at intraseasonal to multi-decadal time scales for North America and adjoining seas. The results show a mixture of performance, with some aspects of climate well reproduced but others poorly so. Although ISV in the eastern Pacific exhibits significant impacts on regional weather and climate, it remains a great challenge for CMIP5 models to realistically represent this form of variability. Among sixteen CMIP5 models examined in this study, only seven capture reasonably well the spatial pattern of the leading ISV mode over the EPAC, although even these models have biases in simulating the ISV amplitude. It is suggested that model fidelity in representing EPAC ISV is closely associated with ability to simulate a realistic summer mean state. The presence of mean westerly or very weak easterly low-level winds over the EPAC warm pool region tends to be conducive for realistic simulations of EPAC ISV. On the other hand, the midsummer drought over Central America is not an enigmatic challenge to all global models, which is similar to CMIP3 results. Out of the 23 CMIP5 models analyzed, roughly half do a reasonable job in simulating the MSD on an individual basis, with a handful performing very well.

Tropical disturbances and cyclone-like storms in the Atlantic and eastern North Pacific were evaluated. When all tropical disturbances and storms are considered, the set

of models examined does a reasonable job of depicting the density and location of activity but with a good deal of scatter across models. In contrast, when only tropical cyclones are analyzed, all models severely under predict the number of TCs, especially in the peak hurricane season. Experiments with the higher resolution GFDL model, indicate that 25-50km resolution models can do a good job at replicating TC frequency and variability, although they still cannot simulate the most intense storms. These biases have implications for simulated precipitation over the southeastern US which can receive up to 20% of annual and 30% of warm season precipitation from TCs (Kam et al., 2012) and contribute significantly to heavy precipitation (Knight and Davis, 2009).

The CMIP5 multi-model ensemble mean reproduces the frequency and mean amplitude of ENSO events, but does not appear to have improved significantly since CMIP3 (Guilyardi et al., 2012). The models do not fully reproduce the phase-locking of ENSO to the seasonal cycle, a deficiency noted in CMIP3 models as well (Guilyardi et al. 2009). Comparison of surface air temperature and precipitation composites associated with ENSO events from 17 CMIP5 models show a wide range in model fidelity to the observed relationships, with the largest discrepancies occurring during winter and spring. Again the situation has not changed since CMIP3 (Joseph and Nigam, 2006; Mo, 2010). The observed response pattern in US winter air temperature for the two types of ENSO (EP and CP) is a warm northeast to cold southwest pattern for the EP El Niño and a warm northwest to cold southeast pattern for the CP El Niño. There are only a few CMIP5 models that can produce the two types of ENSO and simulate the different winter air temperature responses. Overall there has been an improvement since CMIP3 as the CMIP5 models can (1) better simulate the observed spatial patterns of the two types of

779 ENSO and (2) have a significantly smaller inter-model diversity in ENSO intensities
780 (Kim and Yu, 2012). ENSO asymmetry remains a common bias in climate models with
781 implications for simulating tropical decadal variability. Weaker asymmetry in a model
782 indicates that the model is in a different dynamical regime than in the observations (Liang
783 et al., 2012), in other words, the model has a weaker radiative forcing relative to its
784 thermal and mechanical dissipation.

785 Five models were analyzed in terms of circulation anomalies and moisture flows
786 associated with persistent wet and dry summers in the central US. Some of the models
787 have shown improvement, compared to the CMIP3 models, in simulating the major
788 summer circulation components in the Great Plains, e.g., the LLJ, and their seasonal
789 transitions (see Sheffield et al. 2012), a result largely attributable to the higher spatial
790 resolution of CMIP5 models. However, the CMIP5 models still struggle to describe the
791 weather systems that develop around the seasonal circulation and their consequent effects
792 on summer precipitation and extreme hydrological events. This suggests that simulated
793 summer season precipitation anomalies and hydrological extremes in the Great Plains
794 result from processes in which the LLJ plays a less important or different role than that
795 suggested by observations.

796 At decadal scale, the models are evaluated for the PDO and AMO, and their
797 impacts on North American climate. Overall, the CMIP5 models perform well in
798 capturing the PDO influence on North American temperature and on West Coast
799 precipitation in winter. The largest deficiencies appear to lie in the representation of the
800 wintertime precipitation signature over the eastern North America. At first, it may be
801 surprising the CMIP5 models perform much better at capturing the PDO temperature and

precipitation patterns than the ENSO temperature and precipitation patterns. However, this disparity likely owes to the more direct connection between the PDO and North American climate. The PDO index closely tracks the strength of the Aleutian low and the associated Pacific/North American (PNA) teleconnection pattern. Therefore, models that successfully reproduce the structure and variability of the PNA are likely to capture the basic climate features associated with the PDO, although problems with the precipitation patterns are evident. ENSO, in contrast, provides a more indirect connection, where variations in the tropical Pacific ENSO SST patterns, associated convection anomalies, and interaction with the mean state can produce downstream teleconnection patterns that vary substantially from the PNA-like patterns typical of ENSO episodes. Therefore, with ENSO there is larger opportunity for model errors in various components of the phenomenon to result in temperature and precipitation patterns that are distinct from observations.

The representation of AMO variability in the models is reasonable, but they fail to capture the timing of the warm and cool phases during the 20th century (also seen in CMIP3 models; Knight, 2008). They do, however, simulate warming in the North Atlantic in recent decades in line with observations. Teleconnections between the AMO and North American climate are not well represented. Less than half of the 23 models examined are able to replicate the observed relationship between the AMO and vertical wind shear, with implications for their ability to simulate the influence of AMO variability on tropical cyclones in the North Atlantic. Both the inaccurate representations of the AMO itself and model wind response to the AMO in the tropical North Atlantic can result in the model discrepancies in simulating the relationship of the AMO with

tropical cyclones in the North Atlantic. Additionally, given the relatively low resolutions of current CMIP5 models, future modeling work should emphasize to resolve both tropical cyclones and large-scale climate variability and their potential interaction. The SST structure and hydroclimate impact of the AMO represents a real challenge for the CMIP5 models analyzed. While the majority of models have features resembling the observed structure of the SST anomalies they still have room to improve the structure and magnitude of the SST anomalies. Except by one of the models, the buildup of the drying conditions over the central US as well as the wet conditions along the coastal south Atlantic US states from summer to fall are not well captured by the models.

The observed warming hole over the central to southeast US is replicated only by some models indicating that it is driven partly by decadal variability, rather than a forced climate signal or land surface feedback. The observed warming hole in the eastern US is closely associated with the multi-decadal oscillation in North Atlantic (65-70 years cycle; Kumar et al., 2012). The multi-model mean fails to reproduce observed increases in 20th century precipitation, but rather shows a robust drying signal in the southern US and Mexico. The failure is in part because of higher uncertainty among CMIP5 models related to sign and magnitude of precipitation change at regional scale. The drying signal in the models is symptomatic of CMIP3 models also (IPCC, 2007) and is likely driven by the inadequate connection between increasing precipitation and global SST warming, at least for summer, in the majority of models as shown by Fu et al. (2012) for the southern US. The observed change in winter and summer climatological precipitation from the first half (1901-1950) to the second half (1951-1999) of the 20th century is not well simulated by the CMIP5 models. While some success is attained by some models in

simulating the increased precipitation over the Pacific Northwest in winter, the drying of the US coastal plains is only partially captured by some models.

For the climate features and models analyzed here, there does not appear to be a great deal of improvement since CMIP3. For example, CMIP5 models still cannot capture the seasonal timing of ENSO events, that tend to peak in the fall and winter, and the spurious drying signal in the southern US and Mexico continues from CMIP3. However, some features continue to be well simulated, such as the SST pattern of the PDO, and features related to spatial resolution are likely to have improved, such as the representation of TCs and terrain-dependent regional processes. Overall, the models are less able to capture observed variability and long-term trends than they are the mean climate state as evaluated in the first part of this paper (Sheffield et al., 2012). This is understandable for decadal to multi-decadal variability which is dependent on the models' internal variability or the sensitivity to external forcing. Some of the biases, however, are related to problems in simulating the mean state.

Acknowledgements. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The authors acknowledge the support of NOAA Climate Program Office

870 Modeling, Analysis, Predictions and Projections (MAPP) Program as part of the CMIP5
871 Task Force.
872

References

- Aiyyer, A. R., and C. Thorncroft, 2006: Climatology of vertical wind shear over the tropical Atlantic. *J. Clim.*, **19**, 2969-2983.
- Aiyyer, A. and J. Molinari, 2008: MJO and Tropical Cyclogenesis in the Gulf of Mexico and Eastern Pacific: Case Study and Idealized Numerical Modeling. *J. Atmos. Sci.*, **65**, 2691-2704.
- Arora, V. K., J. F. Scinocca, G. J. Boer, J. R. Christian, K. L. Denman, G. M. Flato, V. V. Kharin, W. G. Lee, and W. J. Merryfield, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.
- Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, doi:10.1029/2006JC003798.
- Bao, Q., and co-authors, 2012: The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2. *Adv. Atmos. Sci.*, submitted.
- Barrett, B. S. and L. M. Leslie, 2009: Links between Tropical Cyclone Activity and Madden-Julian Oscillation Phase in the North Atlantic and Northeast Pacific Basins. *Mon. Weather Rev.*, **137**, 727-744. doi:10.1175/2008MWR2602.1
- Bengtsson, L., H. Böttger, and M. Kanamitsu, 1982: Simulation of hurricane-type vortices in a general circulation model. *Tellus*, **34**, 440-457.
- Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **B57**, 289-300.

895 Bengtsson, L., K. Hodges, M. Esch, N. Keenlyside, L. Kornblueh, J.-J. Luo, and T.
896 Yamagata, 2007: How may tropical cyclones change in a warmer climate. *Tellus*,
897 **59A**, 539–561.

898 Bi, D., M. Dix, S. Marsland, T. Hirst, S. O’Farrell and coauthors, 2012: ACCESS: The
899 Australian Coupled Climate Model for IPCC AR5 and CMIP5. AMOS conference,
900 2012, Sydney, Australia (available online at
901 <https://wiki.csiro.au/confluence/display/ACCESS/ACCESS+Publications>)

902 Biasutti, M., and A. H. Sobel, 2009: Delayed Sahel rainfall and global seasonal cycle in a
903 warmer climate. *Geophys. Res. Lett.*, **36**. doi:10.1029/2009GL041303

904 Burgers, G., and D. B. Stephenson, 1999: The “normality” of El Niño. *Geophys. Res.*
905 *Lett.*, **26**, 1027–1030.

906 Camargo, S. J. and S. E. Zebiak, 2002: Improving the detection and tracking of tropical
907 cyclones in atmospheric general circulation models. *Weather and Forecasting*, **17**,
908 1152–1162

909 Camargo, S. J., A. G. Barnston, and S. E. Zebiak, 2005: A statistical assessment of
910 tropical cyclone activity in atmospheric general circulation models. *Tellus*, **57A**, 589–
911 604.

912 Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007: Tropical cyclone
913 genesis potential index in climate models. *Tellus*, **59A**, 428–443.

914 Cavalieri, D., C. Parkinson, P. Gloerson, and H. J. Zwally. 1997, updated 2005. /Sea ice
915 concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data,
916 June to September 2001/. Boulder, CO: National Snow and Ice Data Center. Digital
917 media and CD-ROM.

918 Chylek, P., J. Li, M. K. Dubey, M. Wang, and G. Lesins, 2011: Observed and model
 919 simulated 20th Century Arctic temperature variability: Canadian Earth System Model
 920 CanESM2. *Atmospheric Chemistry and Physics Discussions*, **11** (8), 22,893–22,907,
 921 doi: 10.5194/acpd-11-22893-2011

922 Colle, B. A., Z. Zhang, K. Lombardo, and P. Liu, 2012: Historical and future changes in
 923 western Atlantic extratropical cyclones in the CMIP5 models. *J. Climate*, submitted.

924 Collins, M., S. F. B. Tett, and C. Cooper, 2001: The internal climate variability of
 925 HadCM3, a version of the Hadley Centre Coupled Model without flux adjustments.
 926 *Climate Dynamics*, **17** (1), 61-81.

927 Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E.
 928 Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R. I.
 929 Crouthamel, A. N. Grant, P. Y. Groisman, P. D. Jones, M. Kruk, A. C. Kruger, G. J.
 930 Marshall, M. Maugeri, H. Y. Mok, Ø. Nordli, T. F. Ross, R. M. Trigo, X. L. Wang, S.
 931 D. Woodruff, and S. J. Worley, 2011, The Twentieth Century Reanalysis Project.
 932 *Quarterly J. Roy. Meteorol. Soc.*, **137**, 1-28. DOI: 10.1002/qj.776.

933 Dai, A., K. E. Trenberth, and T. R. Karl, 1999: Effects of clouds, soil moisture,
 934 precipitation, and water vapor on diurnal temperature range. *J. Climate*, **12**, 2451–
 935 2473, doi: 10.1175/1520-0442(1999)0122.0.CO;2.

936 de Szoeke, S. P. and C. S. Bretherton, 2005: Variability in the southerly flow into the
 937 eastern Pacific ITCZ. *J. Atmos. Sci.*, **62**, 4400-4411.

938 Donner, L. J., with 28 co-authors, 2011: The dynamical core, physical parameterizations,
 939 and basic simulation characteristics of the atmospheric component AM3 of the GFDL
 940 Global Coupled Model CM3. *J. Climate*, **24**, doi:10.1175/2011JCLI3955.1.

941 Dufresne, J-L., and 58 co-authors, N, 2012: Climate change projections using the IPSL-
 942 CM5 Earth System Model: from CMIP3 to CMIP5, *Clim. Dyn.*, submitted.
 943 Easterling, D.R., Briony Horton, Philip D. Jones, Thomas C. Peterson, Thomas R. Karl,
 944 David E. Parker, M. James Salinger, Vyacheslav Razuvayev, Neil Plummer, Paul
 945 Jamason and Christopher K. Folland, 1997: *Science*, **277** (5324), 364-367. DOI:
 946 10.1126/science.277.5324.364.
 947 Emanuel, K., R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming:
 948 Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.*, **89**, 347
 949 – 367.
 950 Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal
 951 oscillation and its relationship to rainfall and river flows in the continental US.
 952 *Geophys. Res. Lett.*, **28**, 2077-2080.
 953 Fetterer, F., K. Knowles, W. Meier, and M. Savoie. 2002, updated 2009. /Sea Ice Index/.
 954 Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.
 955 Furtado, J. C., E. Di Lorenzo, N. Schneider, and N. Bond, 2011: North Pacific decadal
 956 variability and climate change in the IPCC AR4 models. *J. Climate*, **24**, 3049-3067.
 957 Gent, P. R., and Coauthors, 2011: The Community Climate System Model Version 4. *J.*
 958 *Climate*, **24**, 4973–4991. doi: <http://dx.doi.org/10.1175/2011JCLI4083.1>
 959 Giorgi, F. and R. Francisco, 2000: Uncertainties in regional climate change prediction: a
 960 regional analysis of ensemble simulations with the HADCM2 coupled AOGCM.
 961 *Climate Dynamics*, **16** (2-3), 169–182, doi:10.1007/PL00013733.

962 Goldenberg, S. B., C. W. Landsea, A. M. Maestas-Nunez, and W. M. Gray, 2001: The
 963 recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**,
 964 474-479.

965 Gualdi, S., E. Scoccimarro, and A. Navarra, 2008: Changes in tropical cyclone activity
 966 due to global warming: results from a high-resolution coupled general circulation
 967 model. *J. Climate*, **21**, 5204–5228.

968 Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van
 969 Oldenborgh, and T. Stockdale, 2009: Understanding El Niño in Ocean–Atmosphere
 970 General Circulation Models: Progress and Challenges. *Bull. Amer. Meteor. Soc.*, **90**,
 971 325–340.

972 Guilyardi, E., H. Bellenger, M. Collins, S. Ferrett, W. Cai, and A. Wittenberg, 2012: A
 973 first look at ENSO in CMIP5. *CLIVAR Exchanges*, **17**, 29-32. ISSN: 1026-0471.

974 Hazeleger, W., and 31 co-authors, 2010: EC-Earth: A seamless Earth system prediction
 975 approach in action. *Bull. Amer. Meteor. Soc.*, **91**, 1357-1363, doi:
 976 10.1175/2010BAMS2877.1

977 Held, I. and M. Zhao, 2011: The response of tropical cyclone statistics to an increase in
 978 CO2 with fixed sea surface temperatures. *J. Climate*, **24**, 5353–5364.

979 Held, I. M. and coauthors, 2012: High resolution AMIP simulation for CMIP5 using
 980 GFDL’s HIRAM model. *J. Climate*, manuscript in preparation.

981 Higgins, R. W., A. Leetmaa, Y. Xue, and A. Barnston, 2000: Dominant factors
 982 influencing the seasonal predictability of U.S. precipitation and surface air
 983 temperature. *J. Clim.*, **13**, 3994-4017.

984 Higgins, R. W. and W. Shi, 2001: Intercomparison of the Principal Modes of Interannual
 985 and Intraseasonal Variability of the North American Monsoon System. *J. Clim.*, **14**,
 986 403-417.

987 Hodges, K. I., 1994: A general method for tracking analysis and its application to
 988 meteorological data. *Mon. Wea. Rev.*, **122**, 2573–2586.

989 Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F.
 990 Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis:
 991 Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale.
 992 *J. Hydrometeor.*, **8** (1), 38–55.

993 IPCC, 2007: Climate Change 2007: Synthesis Report. Contribution of working groups I,
 994 II, and III to the Fourth Assessment Report of the Intergovernmental Panel on
 995 Climate Change [Core Writing Team, Pachauri, R. K., and Reisinger, A (eds.)] IPCC,
 996 Geneva, Switzerland, 104 pp.

997 Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape
 998 for the North Atlantic Basin, 1886-1983: Contents, limitations and uses, Tech. Memo.
 999 NWS NHC-22, 21 pp., NOAA, Silver Spring, Md.

1000 Jiang, X. and D. E. Waliser, 2008: Northward propagation of the subseasonal variability
 1001 over the eastern Pacific warm pool. *Geophys. Res. Lett.*, **35**, L09814,
 1002 doi:10.1029/2008GL033723.

1003 Jiang, X., and D. E. Waliser, 2009: Two dominant subseasonal variability modes of the
 1004 eastern Pacific ITCZ. *Geophys. Res. Lett.*, **36**, L04704, doi: 10.1029/2008gl036820.

1005 Jiang, X., D. Waliser, D. Kim, M. Zhao, K. Sperber, W. Stern, S. Schubert, G. Zhang, W.
 1006 Wang, M. Khairoutdinov, R. Neale, and M.-I. Lee, 2011: Simulation of the

1007 intraseasonal variability over the Eastern Pacific ITCZ in climate models. *Climate*
1008 *Dyn.*, 10.1007/s00382-011-1098-x, 1-20.

1009 Jin, F-F., S-I. An, A. Timmermann, and J. Zhao, 2003: Strong El Niño events and
1010 nonlinear dynamical heating. *Geophys. Res. Lett.*, **30**, 1120.
1011 doi:10.1029/2002GL016356.

1012 Jones, C. D., and others, 2011: The HadGEM2-ES implementation of CMIP5 centennial
1013 simulations, *Geosci. Model Dev.*, **4**, 543-570, doi:10.5194/gmd-4-543-2011.

1014 Joseph, R., and S. Nigam, 2006: ENSO Evolution and Teleconnections in IPCC's 20th
1015 Century Climate Simulations: Realistic Representation? *J. Climate*, **19**, 4360-4377.

1016 Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S.
1017 Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J.
1018 Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetman, R. Reynolds, R. Jenne,
1019 and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Am.*
1020 *Meteorol. Soc.*, **77**, 437–471

1021 Kam, J., J. Sheffield, X. Yuan, and E. F. Wood, 2012: The influence of Atlantic tropical
1022 cyclones on drought over the eastern US (1980-2007), *J. Climate*, in review.

1023 Karl, Thomas R., George Kukla, Joyce Gavin, 1986: Relationship between Decreased
1024 Temperature Range and Precipitation Trends in the United States and Canada, 1941–
1025 80. *Climate Appl. Meteor.*, **25**, 1878–1886.

1026 Karl T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency,
1027 and intensity in the United States. *Bull. Am. Meteorol. Soc.*, **79** (2), 231-241.

1028 Karnauskas, K. B., A. Giannini, R. Seager, and A. J. Busalacchi, 2012: A simple
 1029 mechanism for the climatological midsummer drought along the Pacific coast of
 1030 Central America. *Atmósfera*, in revision.
 1031 Kavvada, A., A. Ruiz-Barradas, and S. Nigam, 2012: AMO's structure and climate
 1032 footprint in observations and IPCC AR5 climate simulations. *Climate Dynamics*,
 1033 submitted.
 1034 Kayano, M. T. and V. E. Kousky, 1999: Intraseasonal (30-60 day) variability in the
 1035 global tropics: principal modes and their evolution. *Tellus Series a-Dynamic*
 1036 *Meteorology and Oceanography*, **51**, 373-386.
 1037 Kim, D., et al., 2009: Application of MJO Simulation Diagnostics to Climate Models. *J.*
 1038 *Clim.*, **22**, Doi 10.1175/2009jcli3063.1, 6413-6436.
 1039 Kim, D., A. H. Sobel, A. D. Del Genio, Y. Chen, S. Camargo, M.-S. Yao, M. Kelley, and
 1040 L. Nazarenko, 2012: The tropical subseasonal variability simulated in the NASA
 1041 GISS general circulation model, *J. Clim.*, in press.
 1042 Kim, S. T. and J.-Y. Yu, 2012: The two types of ENSO in CMIP5 Models, *Geophysical*
 1043 *Research Letters*, **39**, L11704, doi:10.1029/2012GL052006.
 1044 Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010:
 1045 The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying
 1046 tropical cyclone best track data. *Bull. Amer. Meteor. Soc.*, **91**, 363-376.
 1047 doi:10.1175/2009BAMS2755.1
 1048 Knight, D. B., and R. E. Davis, 2009: Contribution of tropical cyclones to extreme
 1049 rainfall events in the southeastern United States. *J. Geophys. Res.-Atmos.*, **114**, 17.

1050 Knight, J. R. et al., 2005: A signature of persistent natural thermohaline circulation cycles
 1051 in observed climate. *Geophys. Res. Lett.*, **32**, doi:10.1029/2005GL024233.

1052 Knutson, T. R. and K. M. Weickmann, 1987: 30-60 Day Atmospheric Oscillations:
 1053 Composite Life Cycles of Convection and Circulation Anomalies. *Mon. Weather*
 1054 *Rev.*, **115**, 1407-1436.

1055 Knutson, T. R., J. J. Sirutis, S. T. Garner, G. A. Vecchi, and I. M. Held, 2008: Simulated
 1056 reduction in Atlantic hurricane frequency under twenty-first-century warming
 1057 conditions. *Nature Geoscience*, **1**, 359-364.

1058 Knutson, T. R., J. L. McBride, J. Chan, K. Emanuel, G. Holland, C. Landsea, I. Held, J.
 1059 P. Kossin, A. K. Srivastava, and M. Sugi, 2010: Tropical cyclones and climate
 1060 change. *Nature Geoscience*, **3**, 157-163.

1061 Kug, J.-S., F.-F. Jin, and S.-I. An, 2009: Two-types of El Niño events: Cold Tongue El
 1062 Niño and Warm Pool El Niño. *J. Climate*, **22**, 1499-1515.

1063 Kumar, S., V. Merwade, J. Kam, and K. Thurner, 2009: Streamflow trends in Indiana:
 1064 effects of long term persistence, precipitation and subsurface drains, *J. Hydrology*,
 1065 **374** (1-2), 171–183

1066 Kumar, S., J. Kinter, P. A. Dirmeyer, Z. Pan, and J. Adams, 2012: Multi-decadal climate
 1067 variability and the "warming hole" in North America - results from CMIP5 climate
 1068 simulations. *J. Climate*, submitted.

1069 Larkin, N. K. and D. E. Harrison, 2005: On the definition of El Niño and associated
 1070 seasonal average U.S. weather anomalies. *Geophys. Res. Lett.*, **32**:L13705.
 1071 doi:10.1029/2005GL022738.

1072 LaRow, T., Y.-K. Lim, D. Shin, E. Chassignet, and S. Cocke, 2008: Atlantic basin
1073 seasonal hurricane simulations. *J. Climate*, **21**, 3191–3206.

1074 Lauritsen, R., and J. Rogers, 2012: U.S. diurnal temperature range variability and
1075 regional causal mechanisms, 1901-2002. *J. Climate*, doi:10.1175/JCLI-D-11-00429.1,
1076 in press.

1077 Lee, T., and M. J. McPhaden, 2010: Increasing intensity of El Niño in the central-
1078 equatorial Pacific. *Geophys. Res. Lett.*, **37**, L14603, doi:10.1029/2010GL044007.

1079 Lettenmaier D. P., E. F. Wood, J. R. Wallis, 1994: Hydro-climatological trends in the
1080 continental United States, 1948-88. *J. Climate*, **7**, 586-607.

1081 Liang, J., X.-Q. Yang, and D.-Z. Sun, 2012: The effect of ENSO events on the Tropical
1082 Pacific Mean Climate: Insights from an Analytical Model. *J. Climate*, Accepted.

1083 Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination
1084 by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59.

1085 Lorenz, D. J. and D. L. Hartmann, 2006: The Effect of the MJO on the North American
1086 Monsoon. *J. Clim.*, **19**, 333-343.

1087 Magaña, V., J.A. Amador, and S. Medina, 1999: The Midsummer Drought over Mexico
1088 and Central America. *J. Climate*, **12**, 1577–1588.

1089 Maloney, E. D. and D. L. Hartmann, 2000a: Modulation of Eastern North Pacific
1090 Hurricanes by the Madden-Julian Oscillation. *J. Clim.*, **13**, 1451-1460.

1091 Maloney, E. D. and D. L. Hartmann, 2000b: Modulation of hurricane activity in the Gulf
1092 of Mexico by the Madden-Julian oscillation. *Science*, **287**, 2002-2004.

1093 Maloney, E. D. and S. K. Esbensen, 2003: The Amplification of East Pacific Madden-
 1094 Julian Oscillation Convection and Wind Anomalies during June-November. *J. Clim.*,
 1095 **16**, 3482-3497.
 1096 Maloney, E. D. and S. K. Esbensen, 2005: A Modeling Study of Summertime East
 1097 Pacific Wind-Induced Ocean-Atmosphere Exchange in the Intraseasonal Oscillation.
 1098 *J. Clim.*, **18**, 568-584.
 1099 Maloney, E. D. and S. K. Esbensen, 2007: Satellite and Buoy Observations of Boreal
 1100 Summer Intraseasonal Variability in the Tropical Northeast Pacific. *Mon. Weather*
 1101 *Rev.*, **135**, 3-19.
 1102 Maloney, E. D., D. B. Chelton, and S. K. Esbensen, 2008: Subseasonal SST variability in
 1103 the tropical eastern north Pacific during boreal summer. *J. Clim.*, **21**, Doi
 1104 10.1175/2007jcli1856.1, 4149-4167.
 1105 Maloney, E. D., S. J. Camargo, E. Chang, B. Colle, R. Fu, K. L. Geilw, Q. Hu, X. Jiang,
 1106 N. Johnson, K. B. Karaukas, J. Kinter, B. Kirtman, S. Kumar, B. Langenbrunner,
 1107 K. Lombardo, L. Long, A. Mariotti, J. E. Meyerson, K. Mo, J. D. Neelin, Z. Pan, R.
 1108 Seager, Y. Serraw, A. Seth, J. Sheffield, J. Thibeault, S.-P. Xie, C. Wang, B. Wyman,
 1109 and M. Zhao, 2011: North American Climate in CMIP5 Experiments: Part III:
 1110 Assessment of 21st Century Projections. *J. Climate*, submitted.
 1111 Manabe, S., J. L. Holloway, and H. M. Stone, 1970: Tropical circulation in a time-
 1112 integration of a global model atmosphere. *J. Atmos. Sci.*, **27**, 580-613.
 1113 Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. Francis, 1997: A Pacific
 1114 interdecadal climate oscillation with impacts on salmon production. *Bull. Amer.*
 1115 *Meteor. Soc.*, **78**, 1069-1079.

1116 Martin, E. R. and C. Schumacher, 2010: Modulation of Caribbean Precipitation by the
 1117 Madden-Julian Oscillation. *J. Clim.*, doi:10.1175/2010JCLI3773.1, in press.

1118 McCabe, G. J., and D. M. Wolock, 2002: A step increase in streamflow in the
 1119 conterminous United States. *Geophys. Res. Lett.*, **29** (24), 2185, doi:10.1029/
 1120 2002GL015999.

1121 McDonald, R., D. Bleaken, D. Creswell, V. Pope, and C. Senior, 2005: Tropical storms:
 1122 representation and diagnosis in climate models and the impact of climate change.
 1123 *Clim. Dyn.*, **25**, 19–36.

1124 Mo, K. C., 2010: Interdecadal Modulation of the Impact of ENSO on Precipitation and
 1125 Temperature over the United States. *J. Climate*, **23**, 3639–3656.

1126 Magana, V., J. A. Amador, and S. Medina, 1999: The midsummer drought over Mexico
 1127 and Central America. *J. Clim.*, **12**, 1577-1588.

1128 Molinari, J. and D. Vollaro, 2000: Planetary- and synoptic-scale influences on eastern
 1129 Pacific tropical cyclogenesis. *Mon. Weather Rev.*, **128**, 3296-3307.

1130 Nigam, S., B. Guan, and A. Ruiz-Barradas, 2011: Key role of the Atlantic Multidecadal
 1131 Oscillation in 20th century drought and wet periods over the Great Plains. *Geophys.*
 1132 *Res. Lett.*, **38**, L16713, doi:10.1029/2011GL048650.

1133 Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006:
 1134 Tropical cyclone climatology in a global-warming climate as simulated in a 20 km
 1135 mesh global atmospheric model: Frequency and wind intensity analysis. *J. Meteor.*
 1136 *Soc. Japan*, **84**, 259–276.

1137 Osgood, D., A. Giannini, J. Hansen, E. Holthaus, A. Ines, Y. Kaheil, K. Karnauskas, M.
 1138 McLaurin, A. Robertson, K. Shirley, and M. Vicarelli, 2009: Designing index-based

1139 weather insurance for farmers in Central America, Final report to the Commodity
 1140 Risk Management Group, ARD, World Bank. Tech. Rep. No. 09–01. International
 1141 Research Institute for Climate and Society, Palisades, NY, USA, 79 pp.
 1142 Oshima, K., and Y. Tanimoto, 2009: An evaluation of reproducibility of the Pacific
 1143 Decadal Oscillation in the CMIP3 simulations. *J. Meteor. Soc. Japan*, **87**, 755–770.
 1144 Portig, W. H., 1961: Some climatological data of Salvador, Central America. *Weather*,
 1145 **16**, 103–112.
 1146 Rayner, R., D. Parker, E. Horton, C. Folland, L. Alexander, and D. Rowel, 2003: Global
 1147 analyses of sea surface temperature, sea ice, and night marine air temperature since
 1148 the late nineteenth century. *J. Geophys. Res.*, **108**, D14,4407,doi:10.1029/2002JD002
 1149 670.
 1150 Rauscher, S. A., F. Giorgi, N. S. Diffenbaugh, and A. Seth, 2008: Extension and
 1151 intensification of the Meso–American mid–summer drought in the 21st century,
 1152 *Climate Dynamics*, **31**(5), 551–571.
 1153 Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K.
 1154 Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D.
 1155 Lohmann, and D. Toll, 2004: The Global Land Data Assimilation System. *Bull.*
 1156 *Amer. Met. Soc.*, 381–394.
 1157 Rodgers, K. B., P. Friederichs, and M. Latif, 2004: Tropical Pacific decadal variability
 1158 and its relation to decadal modulation of ENSO. *J. Climate*, **17**, 3761–3774.
 1159 Sakamoto, T. T., Y. Komuro, T. Nishimura, M. Ishii, H. Tatebe, H. Shiogama, A.
 1160 Hasegawa, T. Toyoda, M. Mori, T. Suzuki, Y. Imada, T. Nozawa, K. Takata, T.
 1161 Mochizuki, K. Ogochi, S. Emori, H. Hasumi and M. Kimoto, 2012: MIROC4h - a

1162 new high-resolution atmosphere-ocean coupled general circulation model. *J. Meteor.*
 1163 *Soc. Japan*, **90** (3), 325-359.

1164 Santer, B. D. *et al.*, 2000: Statistical significance of trends and trend differences in layer-
 1165 average atmospheric temperature time series. *J. Geophys. Res.* **105**, 7337-7356.

1166 Schopf, P. S., and R. J. Burgman, 2006: A simple mechanism for ENSO residuals and
 1167 asymmetry. *J. Climate*, **19**, 3167–3179.

1168 Serra, Y. L., G. N. Kiladis, and K. I. Hodges, 2010: Tracking and Mean Structure of
 1169 Easterly Waves Over the Intra-Americas Sea. *J. Clim.*, **0**,
 1170 doi:10.1175/2010JCLI3223.1.

1171 Seth, A., S. Rauscher, M. Rojas, S. Camargo, A. Giannini, 2011: Enhanced spring
 1172 convective barrier for monsoons in a warmer world? *Climatic Change Let.*, **104**,
 1173 403~H414, doi:10.1007/s10584-010- 9973-8.

1174 Seth, A., M. Rojas, and S. A. Rauscher, 2010: CMIP3 projected changes in the annual
 1175 cycle of the South American monsoon. *Clim. Change*, **98**: 331~H357.
 1176 doi:10.1007/s10584-009-9736-6

1177 Sheffield, J., A. Barrett, B. Colle, R. Fu, K. L. Geil, Q. Hu, J. Kinter, S. Kumar, B.
 1178 Langenbrunner, K. Lombardo, L. N. Long, E. Maloney, A. Mariotti, J. E. Meyerson,
 1179 K. C. Mo, J. D. Neelin, Z. Pan, A. Ruiz-Barradas, Y. L. Serra, A. Seth, J. M.
 1180 Thibeault, and J. C. Stroeve, 2012: North American Climate in CMIP5 Experiments.
 1181 Part I: Evaluation of 20th Century Continental and Regional Climatology. *J. Climate*,
 1182 submitted.

1183 Small, R. J., S. P. De Szoeke, and S.-P. Xie, 2007: The Central American midsummer
 1184 drought: Regional aspects and large-scale forcing. *J. Clim.*, **20**, Doi
 1185 10.1175/Jcli4261.1, 4853-4873.

1186 Smith, T. M., R.W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to
 1187 NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *J*
 1188 *Clim.*, **21**, 2283-2296.

1189 Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to
 1190 NOAA's Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006).
 1191 *J. Climate*, **21**, 2283-2296.

1192 Sugi, M., A. Noda, and N. Sato, 2002: Influence of the global warming on tropical
 1193 cyclone climatology: An experiment with the JMA global model. *J. Meteor. Soc.*
 1194 *Japan*, **80**, 249–272.

1195 Sun, D.-Z., and T. Zhang, 2006: A regulatory effect of ENSO on the time-mean thermal
 1196 stratification of the equatorial upper ocean. *Geophys. Res. Lett.*, **33**, L07710.
 1197 doi:10.1029/2005GL025296.

1198 Sun, D.-Z., 2010: The Diabatic and Nonlinear Aspects of El Nino Southern Oscillation:
 1199 Implications for its Past and Future Behavior. page 79-104. "Climate Dynamics: Why
 1200 Does Climate Vary?". AGU Geophysical Monograph, Edited by D.-Z. Sun and F.
 1201 Bryan, AGU.

1202 Sun, F., and J-Y. Yu, 2009: A 10–15-year modulation cycle of ENSO intensity. *J.*
 1203 *Climate*, **22**, 1718–1735.

1204 Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the
 1205 experiment design, *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-
 1206 00094.1.
 1207 Trenberth, K. E., D. P. Stepaniak, 2001: Indices of El Niño evolution. *J. Climate*, **14**,
 1208 1697-1701.
 1209 Vecchi, G. A., and B. J. Soden, 2007: Increased tropical Atlantic wind shear in model
 1210 projections of global warming. *Geophys. Res. Lett.*, **34**, 10.1029/2006GL028905.
 1211 Vintzileos, A., M. M. Rienecker, M. J. Suarez, S. D. Schubert, and S. K. Miller, 2005:
 1212 Local versus remote wind forcing of the equatorial Pacific surface temperature in July
 1213 2003. *Geophys. Res. Lett.*, **32**, L05702, doi:10.1029/2004gl02197.
 1214 Voldoire, A., and others, 2012: The CNRM-CM5.1 global climate model: Description
 1215 and basic evaluation, *Clim. Dyn.*, doi:10.1007/s00382-011-1259-y, in press.
 1216 Volodin, E. M., N. A. Diansky, and A. V. Gusev, 2010: Simulating Present-Day Climate
 1217 with the INMCM4.0 Coupled Model of the Atmospheric and Oceanic General
 1218 Circulations. *Izvestia, Atmospheric and Oceanic Physics*, **46**, 414-431
 1219 Vose, R. S., D. R. Easterling, and B. Gleason, 2005: Maximum and minimum
 1220 temperature trends for the globe: An update through 2004, *Geophys. Res. Lett.*, **32**,
 1221 L23822, doi:10.1029/2005GL024379.
 1222 Walsh, K., S. Lavender, H. Murakami, E. Scoccimarro, L.-P. Caron and M. Ghantous,
 1223 2010: The Tropical Cyclone Climate Model Intercomparison Project. *Hurricanes and*
 1224 *Climate Change*, **2**, Springer, 1-24.
 1225 Wang, C., and D. B. Enfield, 2001: The tropical Western Hemisphere warm pool.
 1226 *Geophys. Res. Lett.*, **28**, 1635–1638.

1227 Wang, C., D. B. Enfield, S.-K. Lee, and C. W. Landsea, 2006: Influences of the Atlantic
 1228 warm pool on Western Hemisphere summer rainfall and Atlantic hurricanes. *J. Clim.*,
 1229 **19**, 3011-3028.

1230 Wang, C. and R. H. Weisberg, 2000: The 1997-98 El Niño evolution relative to previous
 1231 El Niño events. *J. Climate*, **13**, 488-501.

1232 Watanabe, M., and Coauthors, 2010: Improved Climate Simulation by MIROC5: Mean
 1233 States, Variability, and Climate Sensitivity. *J. Climate*, **23**, 6312–6335.

1234 Wilks, D. S., 2006: On “field significance” and the false discovery rate. *J. Appl.*
 1235 *Meteorol. Climatol.*, **45**, 1181-1189.

1236 Wu, M. L. C., S. D. Schubert, M. J. Suarez, and N. E. Huang, 2009: An Analysis of
 1237 Moisture Fluxes into the Gulf of California. *J. Clim.*, **22**, Doi 10.1175/2008jcli2525.1,
 1238 2216-2239.

1239 Xie, P., and P.A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on
 1240 gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer.*
 1241 *Meteor. Soc.*, **78**, 2539 - 2558.

1242 Xin X., Wu T., Zhang J., 2012: Introductions to the CMIP 5 simulations conducted by the
 1243 BCC climate system model (in Chinese). *Advances in Climate Change Research*.
 1244 submitted.

1245 Yoshimura, J., M. Sugi, and A. Noda, 2006: Influence of greenhouse warming on tropical
 1246 cyclone frequency. *J. Meteor. Soc. Japan*, **84**, 405–428.

1247 Yu, J.-Y. and H.-Y. Kao, 2007: Decadal Changes of ENSO Persistence Barrier in SST
 1248 and Ocean Heat Content Indices: 1958-2001. *J. Geophys. Res.*, **112**, D13106,
 1249 doi:10.1029/2006JD007654.

1250 Yu., J.-Y. and S. T. Kim, 2010: Identification of Central-Pacific and Eastern-Pacific
 1251 Types of ENSO in CMIP3 Models, *Geophys. Res. Lett.*, **37**, L15705,
 1252 doi:10.1029/2010GL044082.

1253 Yu., J.-Y., Y. Zou, S. T. Kim, and T. Lee, 2012: The changing impact of El Niño on US
 1254 winter temperatures, *Geophys. Res. Lett.*, in press.

1255 Yukimoto, S., et al., 2012: A new global climate model of the Meteorological Research
 1256 Institute: MRI-CGCM3—Model description and basic performance, *J. Meteorol. Soc.
 1257 Jpn.*, **90a**, 23–64.

1258 Zanchettin, D., A. Rubino, D. Matei, O. Bothe, and J. H. Jungclaus, 2012: Multidecadal-
 1259 to-centennial SST variability in the MPI-ESM simulation ensemble for the last
 1260 millennium. *Clim. Dyn.*, **39**, 419-444 doi:10.1007/s00382-012-1361-9.

1261 Zhang, T., D.-Z. Sun, R. Neale, and P. J. Rasch, 2009: An Evaluation of ENSO
 1262 Asymmetry in the Community Climate System Models: A View from the Subsurface.
 1263 *J. Climate*, **22**, 5933-5961.

1264 Zhang, T., and D.-Z. Sun, 2012: An Evaluation of ENSO Asymmetry in CMIP5 Models.
 1265 *J. Climate* (in preparation)

1266 Zhang, Y., J. M. Wallace, and D. S. Battisti, 1997: ENSO-like interdecadal variability:
 1267 1900-93. *J. Climate*, **10**, 1004-1020.

1268 Zhang X., L. A. Vincent, W. D. Hogg, and A. Nittsoo (2000), Temperature and
 1269 precipitation trends in Canada during 20th century, *Atmosphere-Ocean*, **38** (3), 395-
 1270 429.

1271 Zhang, Z. S., Nisancioglu, K., Bentsen, M., Tjiputra, J., Bethke, I., Yan, Q.,
 1272 Risebrobakken, B., Andersson, C., and Jansen, E., 2012: Pre-industrial and mid-

1273 Pliocene simulations with NorESM-L. *Geosci. Model Dev.*, **5**, 523-533,
1274 doi:10.5194/gmd-5-523-2012.

1275 Zhao, M. and I. M. Held, 2012: TC-permitting GCM simulations of hurricane frequency
1276 response to sea surface temperature anomalies projected for the late 21st century. *J.*
1277 *Climate*, **25**, 2995–3009.

1278 Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane
1279 climatology, interannual variability, and response to global warming using a 50km
1280 resolution GCM. *J. Climate*, **22**, 6653–6678.

1281 Zhao, M., I. M. Held, and G. A. Vecchi, 2010: Retrospective forecasts of the hurricane
1282 season using a global atmospheric model assuming persistence of SST anomalies.
1283 *Mon. Wea. Rev.*, **138**, 3858–3868.

1284

1285

1286

Figure Captions

Figure 1. Taylor diagram for summer mean (May-September) rainfall over the eastern Pacific (150°W - 80°W ; 5°S - 30°N) simulated in CMIP5 GCMs. The rainfall observations are based on the Tropical Rainfall Measurement Mission (TRMM).

Figure 2. Spatial distribution of amplitude (a) and phase (b) of the first leading complex EOF (CEOF1) mode based on 30-90-day band-pass filtered TRMM rainfall during boreal summer (June-September) over the eastern Pacific. To make the spatial phase patterns of the CEOF1 based on the observations and simulations comparable to each other, the spatial phase of CEOF1 for each dataset is adjusted by setting the domain averaged value to be zero over a small box region of 110°W - 100°W , 10 - 15°N . Contours are only displayed where the local variance explained by CEOF1 exceeds 8%; (c): X-axis: Pattern correlation coefficients of the CEOF1 mode between TRMM observations and CMIP5 GCM simulations. Y-axis: Relative amplitudes of CEOF1 in model simulations to their observed counterparts. Both pattern correlations and amplitudes are derived by averaging over the area of 5°N - 25°N , 140°W - 80°W where the active ISV is observed. The black “star” mark represents the TRMM observations.

Figure 3. CMIP5 MMM summertime (June-September) precipitation (mm/day) (top) and MSD strength (mm/day) (bottom) from 23 CMIP5 models, historical experiment (1860-2005). Also shown are observed estimates from TRMM 3B43, GPCP, and UNAM. Note that this is the mean of each model’s MSD, not the MSD calculated from the MMM

precipitation climatology. All model output and observational data were regridded onto a common 0.5° grid. The models included in the MMM are: bcc-csm1-1, CanCM4, CanESM2, CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadCM3, HadGEM2-CC, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M.

Figure 4. Storm track density (top) and mean strength (bottom) for ERA Interim and seven CMIP5 models (CanESM2, CCSM4, GFDL-ESM2M, HadGEM2-ES, MIROC5, MPI-ESM-L and MRI-CGCM3). Tracks are based on 6-hourly 850 hPa relative vorticity smoothed to T42 spatial resolution to better capture the synoptic features of the vorticity field.

Figure 5. Tracks of tropical cyclone-like storms in the CMIP5 historical runs in the period 1950-2005 (GFDL-ESM2M (1 ensemble member), HadGEM2 (1), MPI-ESM-LR (3), MRI CGCM3 (5) and MIROC5 (1) models) and in observations for the same period. The number of storms in each case is given in the bottom right corner of each panel. One ensemble member is used for each model.

Figure 6. Mean number of TCs per month in models (GFDL, HadGEM2 (in the figure HGEM2), MPI, MRI, MIROC5) and observations in the North Atlantic (top left panel) and eastern North Pacific (top right panel), using only ensemble 1 for MRI. Number of TCs per year in the period 1950-2005 in models and observations for the North Atlantic

(bottom left panel) and eastern North Pacific (bottom right panel). The blue box shows the 25-75 percentiles range, with the median shown as a red line. The whiskers and red crosses show the data outside of middle quartiles.

Figure 7. A comparison of observed (upper panel) and C180HIRAM simulated (lower panel) hurricane tracks for the N. Atlantic and E. Pacific for 1981-2008. Observations are from the International Best Track Archive for Climate Stewardship (IBTrACS). Only one realization from the 3-member ensemble is shown.

Figure 8. Upper panel: a comparison of observed and C180HIRAM simulated annual hurricane count statistics. Blue boxes show the 25-75 percentiles range, with the median shown as a red line and the mean shown as a red star. The whiskers show the maximum and minimum values. The annual statistics are computed based on 3-member ensemble mean for the 1981-2008 period. Lower panels: Observed and model simulated seasonal cycle (number of hurricanes per month) for the N. Atlantic and E. Pacific from the 3-member ensemble mean (1=JAN, 12=DEC).

Figure 9. Comparison of ENSO-related SAT and precipitation composite patterns between CMIP5 models and observations (CRU TS3.1). Global significance of the differences between SAT (left columns) and precipitation (right columns) composites from CMIP5 models (rows) and observations for each season (designated above each column) and for (a) El Niño and (b) La Niña episodes. The model names are given at the left, and the bottom row corresponds to the CMIP5 ensemble. Light, medium, and dark

red correspond with 90%, 95%, and 99% significance levels, respectively, and blue corresponds to statistically insignificant ($\alpha_{global} > 0.1$) differences. The temperature and precipitation data are regridded to a common equal-area grid with 2° latitude spacing and variable longitude spacing. We determine whether model patterns are significantly different from the observed patterns if the patterns are statistically distinguishable in a test of field significance based on the “false discovery rate” (FDR) (Benjamini and Hochberg 1995, Wilks 2006).

Figure 10. Wintertime composites of ENSO-related SAT and precipitation anomalies in observations and the CMIP5 ensemble. (a-d) Observed and (e-h) CMIP5 ensemble normalized composites of (a,c,e,g) SAT and (b,d,f,h) precipitation in DJF(0/1) for (a,b,e,f) El Niño and (c,d,g,h) La Niña episodes. The contour interval is 0.1°C °C⁻¹ for the SAT composites and 0.05 mm day⁻¹ °C⁻¹ for the precipitation composites, with the zero contour line omitted. In the observed composites, stippling indicates values that are statistically significant at the 95% confidence level based on a two-sided t-test.

Figure 11. US winter surface air temperature regressed on the EP (top) and CP (bottom) ENSO indices from the CMIP5 models and observations (air temperature data are from the NCEP-NCAR reanalysis and SSTs are from the HadISST dataset).

Figure 12. Scatter plot of pattern correlations between the regression patterns from the CMIP5 models and those from the observations (NCEP-NCAR reanalysis and HadISST dataset) for EP versus CP ENSO.

1379

1380 **Figure 13.** The sum of the composite SST anomalies between the two phases of ENSO
1381 from the HadISST observations and CMIP5 coupled models. Presented here are the
1382 results from the coupled models whose corresponding AMIP runs are available for the
1383 analysis. The definition of the warm phase and cold phase of ENSO follows that of
1384 Zhang et al. (2009). The length of data used in the calculation is 50 years for all the
1385 models and observations (1950-99).

1386

1387 **Figure 14.** Standard deviation (upper) and skewness (bottom) of the interannual
1388 variability in Niño-3 SST. Data used are the same as for Figure 13.

1389

1390 **Figure 15.** Reanalysis results: a) and b) show, respectively, summertime precipitation
1391 anomalies (contours) in wet and dry years, in reference to the Great Plains precipitation,
1392 and the vertically integrated moisture fluxes from the surface to the top of the troposphere
1393 (arrows). c) The differences between a) and b). d) and e) show the corresponding 850hPa
1394 geopotential height (countour) and 925hPa wind anomalies (arrows) for the wet and dry
1395 summer, respectively. Their differences are summerized in f).

1396

1397 **Figure 16.** Same as Fig. 15 but for CCSM4 simulation results.

1398

1399 **Figure 17.** PDO SST patterns in observations and CMIP5 models. Linear regression of
1400 SST upon the PDO index in (a) observations and (b) the CMIP5 ensemble, and (c) the
1401 CMIP5 minus observed PDO regression. The contour interval is 0.2°C in (a) and (b) and

0.1°C in (c), with the zero contour omitted. Stippling in (c) indicates where the differences are statistically significant at the 95% confidence level based on a two-sided t-test. (d) Standard deviation of the PDO SST regressions within the ensemble. Contour interval is 0.05°C.

Figure 18. Comparison of PDO SAT and precipitation regression patterns between CMIP5 models and observations. Global significance of the differences between SAT (left columns) and precipitation (right columns) regressions from CMIP5 models (rows) and observations for each season. The model names are given at the left, and the bottom row corresponds to the CMIP5 ensemble. Light, medium, and dark red correspond with 90%, 95%, and 99% significance levels, respectively, and blue corresponds to statistically insignificant ($\alpha_{global} > 0.1$) differences. To highlight multi-decadal variability, a Butterworth 10-year low-pass filter is applied to each PDO index time series, which is then re-standardized. The SAT and precipitation anomalies are regressed in the North America region on this index for each season. The land temperature and precipitation datasets are from CRU TS3.1 for the period of 1901-2009 and all observation and model temperature and precipitation data are regridded to a common equal-area grid, as described in section 5.1. Prior to the regressions, the seasonal SAT and precipitation trend is removed, where the trend is obtained by a LOWESS smoothing with a span of 60 years. Significance is calculated on the regression at each grid point with a two-sided t-test, with the effective degrees of freedom adjusted for the lag-1 autocorrelation in the residuals (Santer et al., 2000).

Figure 19. December-February PDO SAT and precipitation regression patterns over North America. Regressions of DJF SAT (a,c) and precipitation (b,d) upon the PDO index in (a,b) observations and (c,d) the CMIP5 ensemble. The differences between the regression patterns (CMIP5 minus observations) are shown in (e) and (f). The contour interval is 0.1°C for the SAT regressions (a, c, e) and 0.05 mm/day for the precipitation regressions (b,d,f). Stippling in (e) and (f) correspond to differences that are significantly different at the 95% confidence level based on a two-sided t-test.

Figure 20. Time series of the AMO indices from observation and models. The AMO index defined as the detrended North Atlantic SST during the Atlantic hurricane season of June to November (JJASON) from the equator to 60°N, 75°W-5°W with the 11-year running mean. The observed AMO index is plotted in the first panel. The autocorrelations of all AMO indices are plotted in the last panel. SST is from the NOAA extended reconstructed SST version 3 (Smith et al. 2008).

Figure 21. Regressions of the JJASON AMO on vertical wind shear. The observed regression is plotted in the first panel. The vertical wind shear is calculated as the magnitude of the vector difference between winds at 200 mb and 850 mb. The atmospheric reanalysis data set is the NOAA Earth System Research Laboratory (ESRL) 20th Century Reanalysis (20CR) (Compo et al. 2011).

Figure 22. Summer and fall regressions of AMO indices on SST and precipitation in observations from HadISSTv1.1 and CRUTS3.1 data sets, and historical simulations of

the 20th century climate from CMIP5 models for the 1901-1999 common period. AMO indices are calculated as area-averaged SST anomalies over the domain (75°-5°W, 0°-60°N) which are detrended first and then smoothed via a 11-year running mean. Regressions are calculated for each individual ensemble of a given model, then the mean of the regressions is displayed. The number in parenthesis denotes the number of ensembles used from each model. Blue/red shading denotes negative/positive SST anomalies, while brown/green shading denotes negative/positive precipitation anomalies. Contour interval is 0.1K and 0.02 mm day⁻¹.

Figure 23. Observed and multimodel average temperature trends in North America (1930-2004). Observations are from the CRU TS3.1 dataset. The multimodel average is calculated from the best ensemble member from each model, selected based on the highest spatial correlation with observed temperature trends over North America in the respective model. Number in each panel show global mean land only temperature trend (60S to 60N). The models are CanESM2, CCM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadCM3, HadGEM2-CC, HadGEM2-ES, INMCM4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM, MIROC5, MPI-ESM-LR, MRI-CGCM3.

Figure 24. Summer and winter temperature trend (1930-2004), unit: degree C/decade. Multi-model average is based on 19 models, best ensemble member from each model as in Fig. 23. Number in each panel show global mean land only temperature trend (60S to 60N). East and West US regions are also shown by boxes.

1471

1472 **Figure 25.** (a) 30 years running trend in East USA (30N to 47N, and 260E to 280E) DJF
1473 temperature trend (b) 30 year running trend difference between East and West USA.
1474 West USA is defined as (30N to 47N, and 240E to 260E). Shaded region show 95%
1475 uncertainty range calculated from total 96 ensemble members. X axis represent starting of
1476 30 year period, for example trend corresponding to 1930 represent trend from 1930 to
1477 1959. Five core models are also shown (only first ensemble member from each model).

1478

1479 **Figure 26.** Comparison of observed (left) and model simulated linear trends in daily
1480 temperature range ($T_{max} - T_{min}$) during 1951-2000 period. The model ensemble mean
1481 daily temperature range (DTR) is computed from 23 models totaling 109 members
1482 available in the historical experiment. The contour lines are the inter-model variance of
1483 the trends.

1484

1485 **Figure 27.** Observed and multi-model average annual precipitation trend over North
1486 America. Observation is the CRU TS3.1 dataset. The multi-model average is calculated
1487 from 19 climate models, one ensemble member from each model (see Figure 23).

1488

1489 **Figure 28.** (a) 30 years running annual precipitation trend in East USA; (b) 30 year
1490 running annual precipitation trend in West USA. East and West USA regions are shown
1491 in Fig. 27. Shaded region is the 95% uncertainty range calculated from a total of 96
1492 ensemble members. X axis represents the start year of the 30 year period. For example,

the value at 1930 represents the trend from 1930 to 1959. Five example models are also shown (first ensemble member).

Figure 29. Differences between climatologies of the second half (1951-1999) and the first half (1901-1950) of winter and summer precipitation in observations from CRUTS3.1 data set, and historical simulations of the 20th century climate from CMIP5 models. The number in parenthesis denotes the number of ensembles used from each model. Brown/green shading denotes negative/positive precipitation anomalies; contour interval is 0.1 mm day⁻¹.

1503

1504

1504 **Tables**

1505 Table 1. CMIP5 models evaluated and their attributes.

Model	Center	Atmospheric Horizontal Resolution (lon. x lat.)	Number of model levels	Reference
ACCESS1-0	Commonwealth Scientific and Industrial Research Organization/Bureau of Meteorology, Australia	1.875 x 1.25	38	Bi et al. (2012)
BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration, China	2.8 x 2.8	26	Xin et al. (2012)
CanCM4	Canadian Centre for Climate Modelling and Analysis, Canada	2.8 x 2.8	35	Chylek et al. (2011)
CanESM2	Canadian Center for Climate Modeling and Analysis, Canada	2.8 x 2.8	35	Arora et al. (2011)
CCSM4	National Center for Atmospheric Research, USA	1.25 x 0.94	26	Gent et al. (2011)
CESM1-	Community Earth System Model	1.4 x 1.4	26	Gent et al.

CAM5-1-FV2	Contributors (NSF-DOE-NCAR)			(2011)
CNRM-CM5.1	National Centre for Meteorological Research, France	1.4 x 1.4	31	Voldoire et al. (2011)
CSIRO-MK3.6	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, AUS	1.8 x 1.8	18	Rotstayn et al. (2010)
EC-EARTH	EC-EARTH consortium	1.125 x 1.12	62	Hazeleger et al. (2010)
FGOALS-S2.0	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences	2.8 x 1.6	26	Bao et al. (2012)
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 x 2.0	48	Donner et al. (2011)
GFDL-ESM2G/M	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5 x 2.0	48	Donner et al. (2011)

GISS-E2-H/R	NASA Goddard Institute for Space Studies, USA	2.5 x 2.0	40	Kim et al. (2012)
HadCM3	Met Office Hadley Centre, UK	3.75 x 2.5	19	Collins et al. (2001)
HADGEM2- CC (Chemistry coupled)	Met Office Hadley Centre, UK	1.875 x 1.25	60	Jones et al. (2011)
HadGEM2- ES	Met Office Hadley Centre, UK	1.875 x 1.25	60	Jones et al. (2011)
INMCM4	Institute for Numerical Mathematics, Russia	2 x 1.5	21	Volodin et al. (2010)
IPSL-CM5A- LR	Institut Pierre Simon Laplace, France	3.75 x 1.8	39	Dufresne et al. (2012)
IPSL-CM5A- MR	Institut Pierre Simon Laplace, France	2.5 x 1.25	39	Dufresne et al. (2012)
MIROC4h	Atmosphere and Ocean Research Institute (The University of	0.56 x 0.56	56	Sakamoto et al.

	Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan			(2012)
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	1.4 x 1.4	40	Watanabe et al. (2010)
MIROC- ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 x 2.8	80	Watanabe et al. (2010)
MIROC- ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8 x 2.8	80	Watanabe et al. (2010)

MPI-ESM- LR	Max Planck Institute for Meteorology, Germany	1.9 x 1.9	47	Zanchettin et al. (2012)
MRI-CGCM3	Meteorological Research Institute, Japan	1.1 x 1.1	48	Yukimoto et al. (2011)
NorESM1-M	Norwegian Climate Center, Norway	2.5 x 1.9	26	Zhang et al. (2012)

1506

1507

1507 Table 2. Observational and reanalysis datasets used in the evaluations

Dataset	Type	Spatial Domain	Temporal Domain	Reference
<i>Precipitation</i>				
TMPA 3B42 V6	Satellite	0.25 deg, 50S-50N	3-hourly/monthly, 1998-2010	Huffman <i>et al.</i> 2007
CRU TS3.1	Gauge	0.5 deg, global land	Monthly, 1901-2008	Mitchell <i>et al.</i> (2005)
CPC unified	Gauge	0.5 deg, US	Daily, 1948-2010	Xie <i>et al.</i> , 2010
<i>Temperature</i>				
CRU TS3.1	Gauge	0.5 deg, global land	Monthly, 1901-2008	Mitchell <i>et al.</i> (2005)
<i>Sea Surface Temperature</i>				
HadISST	In situ/satellite	Global Oceans, 1.0 deg	Monthly, 1870-present	Rayner <i>et al.</i> (2003)
ERSSTv3b	In situ	Global Oceans, 2.0	Monthly, 1854-present	Smith <i>et al.</i> (2008)

deg

Reanalyses

NCEP-NCAR	Model	~1.9	deg, 6-hourly, 1948-present	Kalnay et al. (1996)
	reanalysis	global		
NCEP-DOE	Model	~1.9	deg, 6-hourly, 1979-present	Kanamitsu et al. (200X)
	reanalysis	global		
20CR	Model	~0.3	deg, 6-hourly, 1871-present	Compo et al. (2011)
	reanalysis	global		
ERA-interim	Model	1.5	deg, 6-hourly, 1979-present	Dee et al. (2011)
	reanalysis	global		

Storm Tracks

National Hurricane Center	In-situ/satellite	Storm tracks, 6-hourly, 1851-present	Jarvinen et al. (1984)
best track tropical cyclone data		Eastern N. (Atlantic), 1949-present (eastern N. Atlantic Pacific and present (eastern N. Atlantic Pacific)	
IBTrACS	Best track datasets from various	Storm 6-hourly, 1842-2010 tracks, global	Knapp et al. (2010)

agencies

oceans

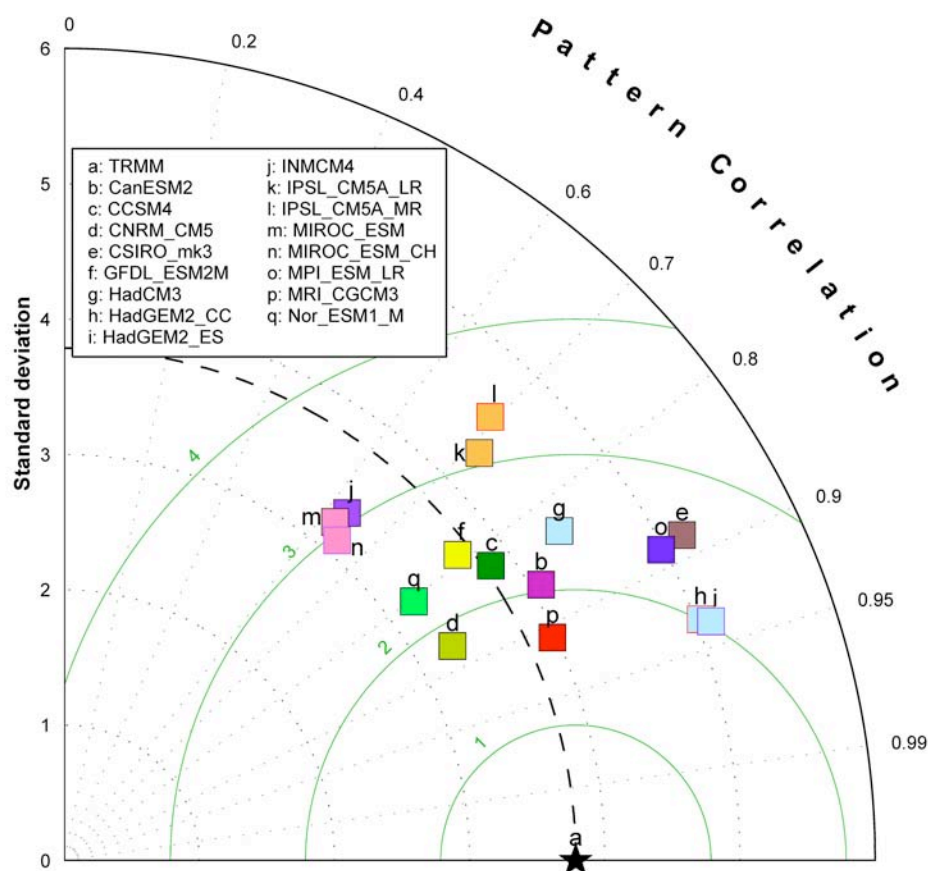
1508

1509

1510

1511

1511



1512

1513

1514

1515

Figure 1. Taylor diagram for summer mean (May-September) rainfall over the eastern Pacific (150°W-80°W; 5°S-30°N) simulated in CMIP5 GCMs. The rainfall observations are based on the Tropical Rainfall Measurement Mission (TRMM).

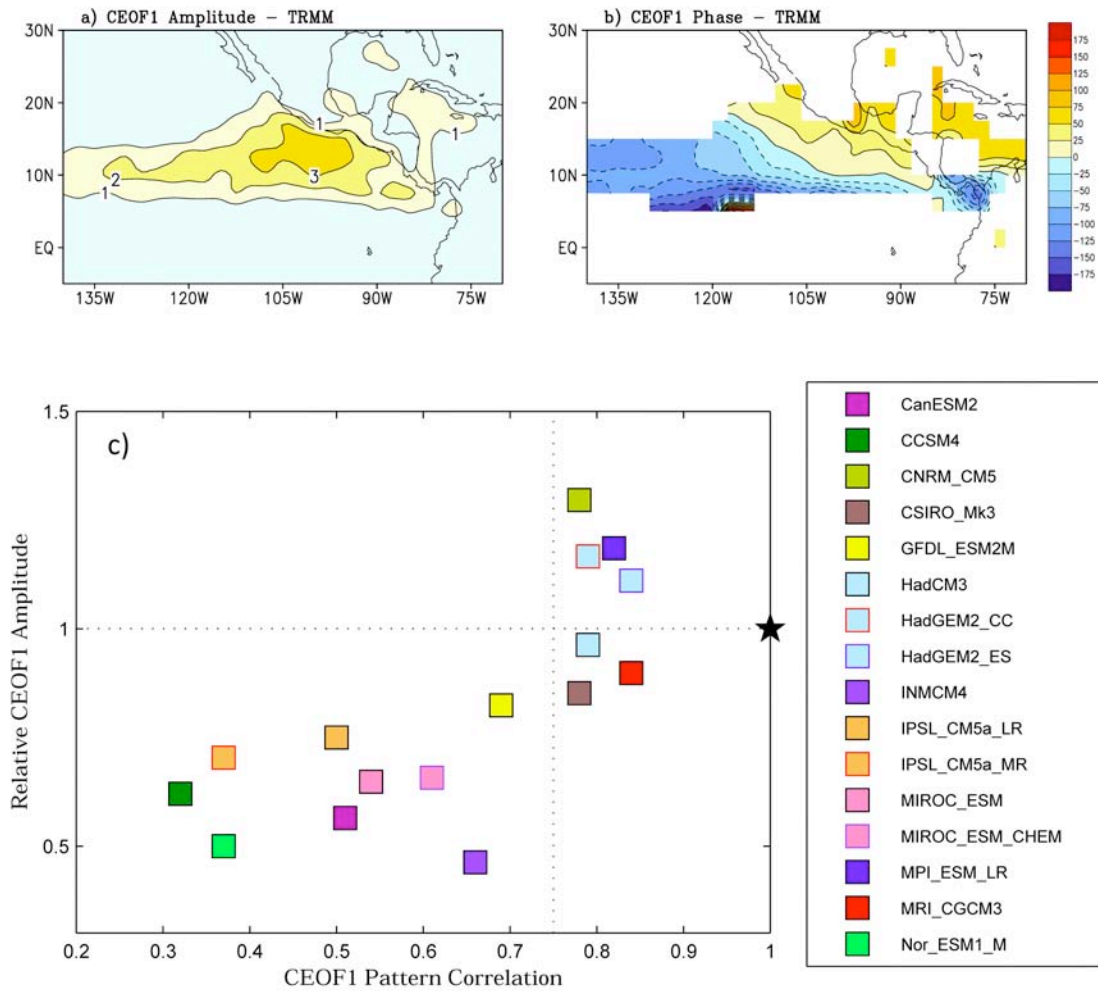
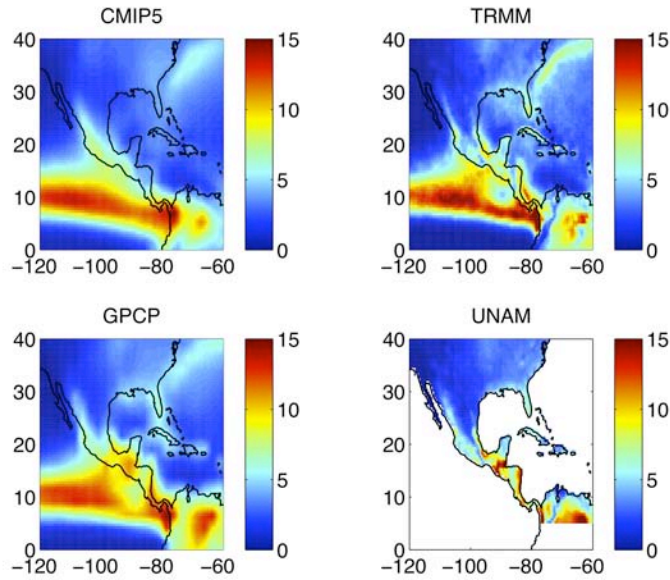
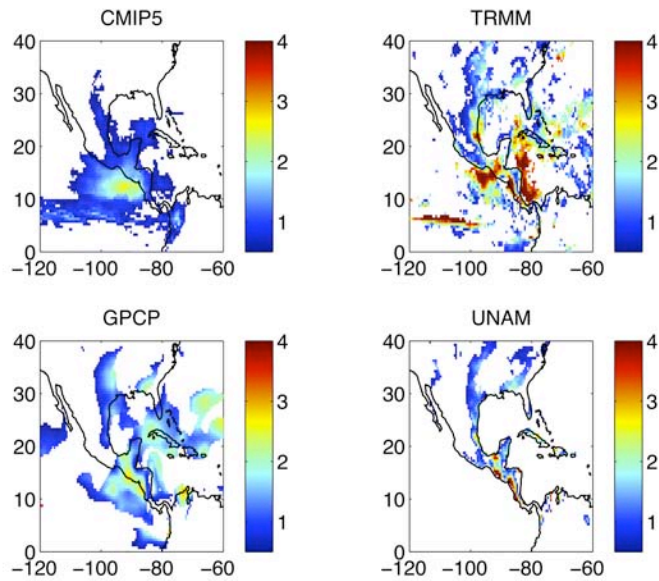


Figure 2. Spatial distribution of amplitude (a) and phase (b) of the first leading complex EOF (CEOF1) mode based on 30-90-day band-pass filtered TRMM rainfall during boreal summer (June-September) over the eastern Pacific. To make the spatial phase patterns of the CEOF1 based on the observations and simulations comparable to each other, the spatial phase of CEOF1 for each dataset is adjusted by setting the domain averaged value to be zero over a small box region of 110°W-100°W, 10-15°N. Contours are only displayed where the local variance explained by CEOF1 exceeds 8%; (c): X-axis: Pattern correlation coefficients of the CEOF1 mode between TRMM observations and CMIP5 GCM simulations. Y-axis: Relative amplitudes of CEOF1 in model simulations to their observed counterparts. Both pattern correlations and amplitudes are derived by averaging over the area of 5°N-25°N, 140°W-80°W where the active ISV is observed. The black “star” mark represents the TRMM observations.



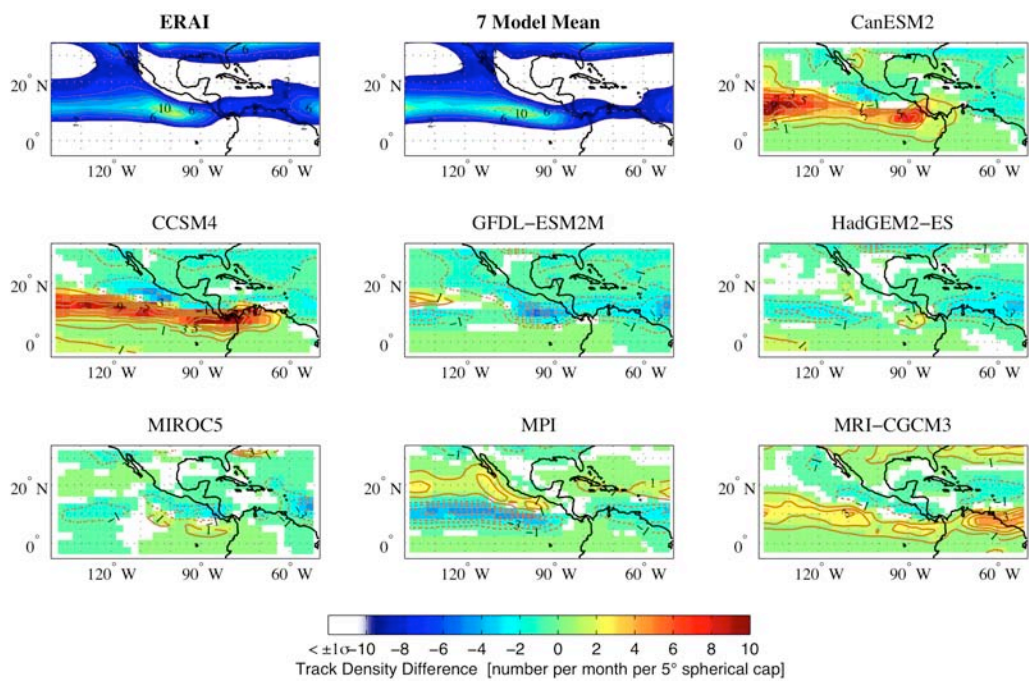
1529



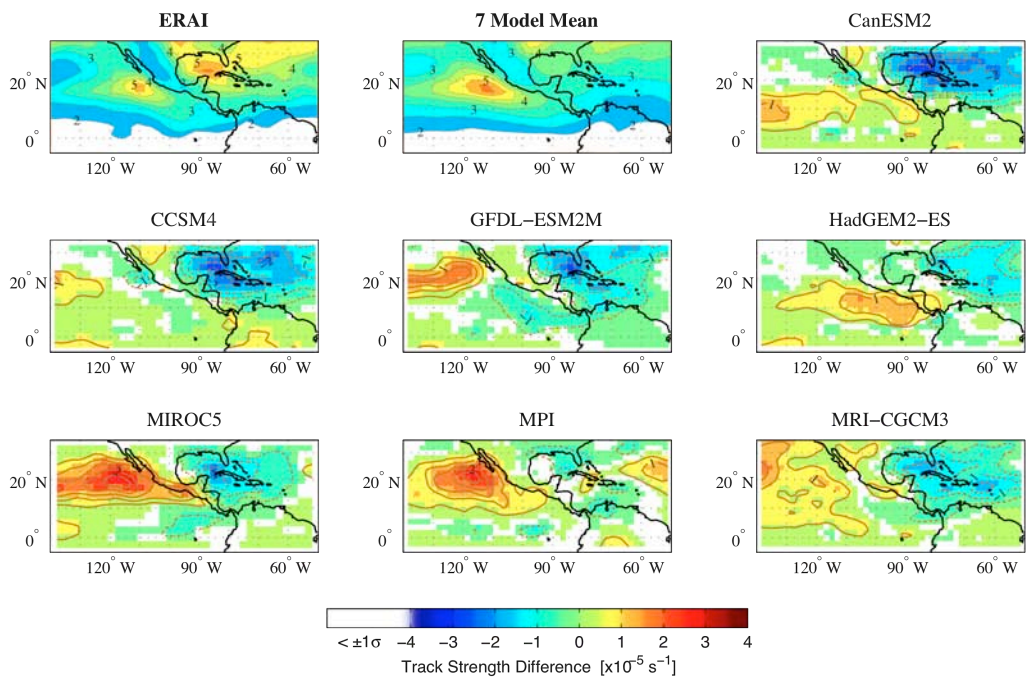
1530

1531 Figure 3. CMIP5 M3M summertime (June-September) precipitation (mm/day) (top) and
 1532 MSD strength (mm/day) (bottom) from 23 CMIP5 models, historical experiment (1860-
 1533 2005). Also shown are observed estimates from TRMM 3B43, GPCP, and UNAM. Note
 1534 that this is the mean of each model's MSD, not the MSD calculated from the M3M
 1535 precipitation climatology. All model output and observational data were regridded onto a
 1536 common 0.5° grid. The models included in the M3M are: bcc-csm1-1, CanCM4,
 1537 CanESM2, CCSM4, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G,
 1538 GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadCM3, HadGEM2-CC, HadGEM2-ES,
 1539 inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MIROC-ESM-
 1540 CHEM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M.
 1541

1541

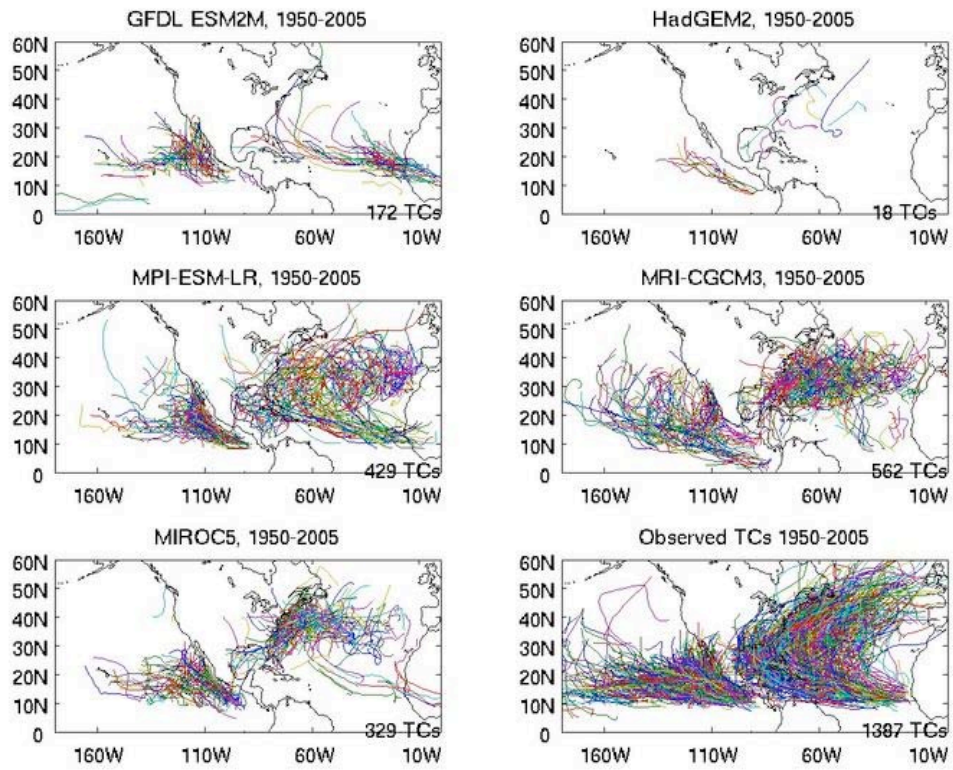


1542



1543

1544 Figure 4. Storm track density (top) and mean strength (bottom) for ERA Interim and
1545 seven CMIP5 models (CanESM2, CCSM4, GFDL-ESM2M, HadGEM2-ES, MIROC5,
1546 MPI-ESM-L and MRI-CGCM3). Tracks are based on 6-hourly 850 hPa relative vorticity
1547 smoothed to T42 spatial resolution to better capture the synoptic features of the vorticity
1548 field.



1549

1550 Figure 5. Tracks of tropical cyclone-like storms in the CMIP5 historical runs in the
 1551 period 1950-2005 (GFDL-ESM2M (1 ensemble member), HadGEM2 (1), MPI-ESM-LR
 1552 (3), MRI CGCM3 (5) and MIROC5 (1) models) and in observations for the same period.
 1553 The number of storms in each case is given in the bottom right corner of each panel. One
 1554 ensemble member is used for each model.

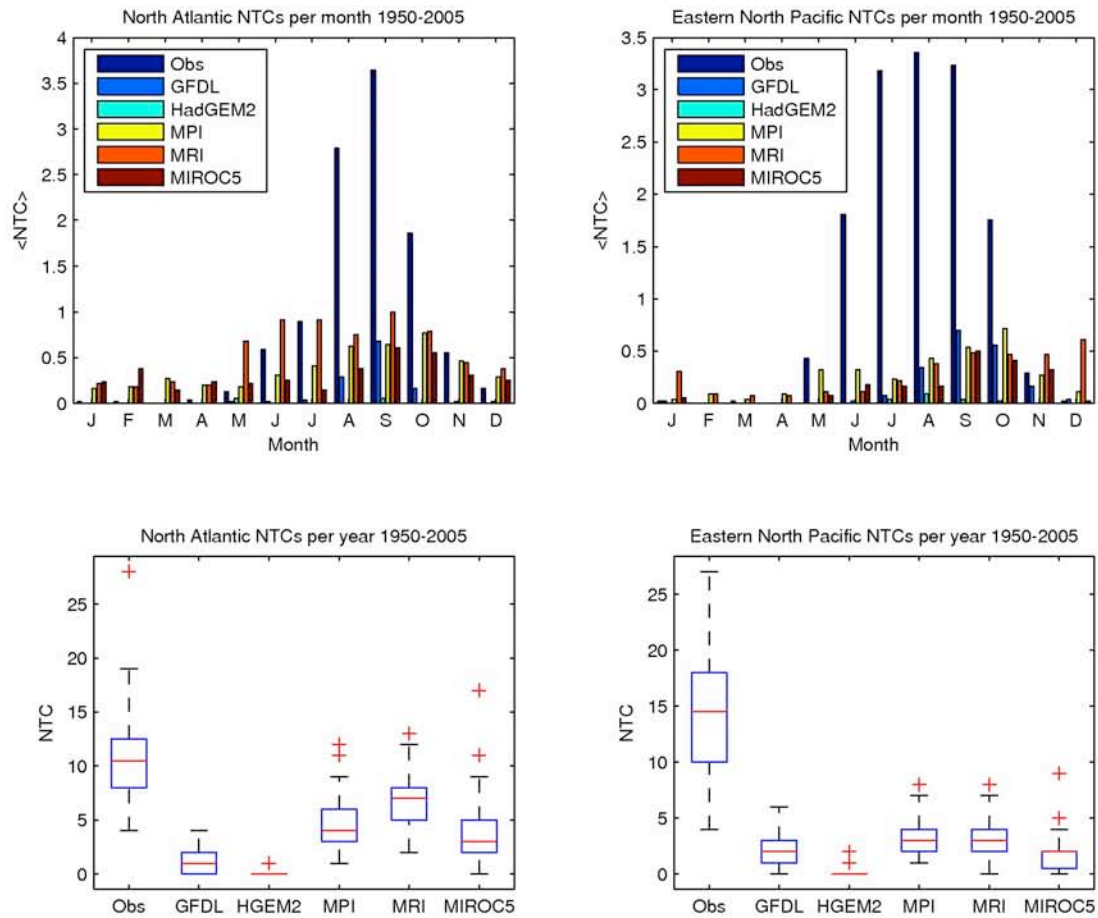


Figure 6. Mean number of TCs per month in models (GFDL, HadGEM2 (in the figure HGEM2), MPI, MRI, MIROC5) and observations in the North Atlantic (top left panel) and eastern North Pacific (top right panel), using only ensemble 1 for MRI. Number of TCs per year in the period 1950-2005 in models and observations for the North Atlantic (bottom left panel) and eastern North Pacific (bottom right panel). The blue box shows the 25-75 percentiles range, with the median shown as a red line. The whiskers and red crosses show the data outside of middle quartiles.

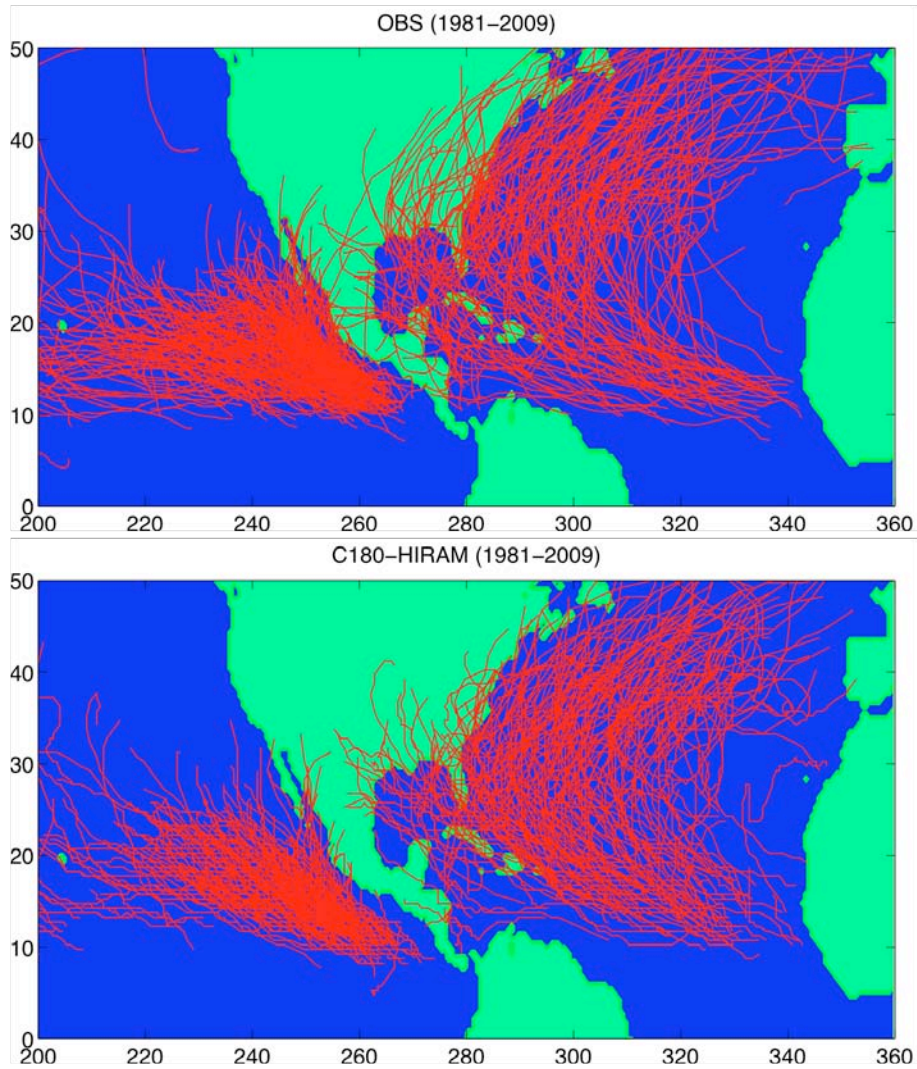


Figure 7. A comparison of observed (upper panel) and C180HIRAM simulated (lower panel) hurricane tracks for the N. Atlantic and E. Pacific for 1981-2008. Observations are from the International Best Track Archive for Climate Stewardship (IBTrACS). Only one realization from the 3-member ensemble is shown.

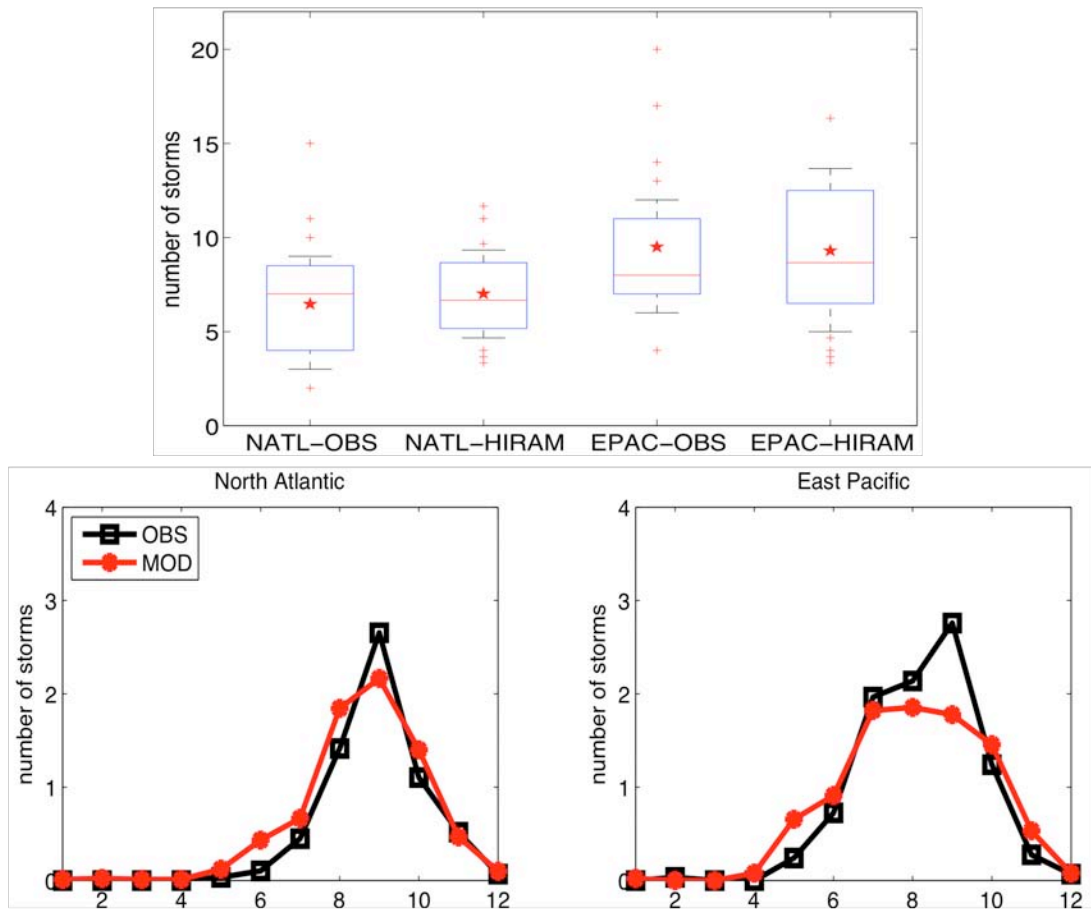


Figure 8. Upper panel: a comparison of observed and C180HIRAM simulated annual hurricane count statistics. Blue boxes show the 25-75 percentiles range, with the median shown as a red line and the mean shown as a red star. The whiskers show the maximum and minimum values. The annual statistics are computed based on 3-member ensemble mean for the 1981-2008 period. Lower panels: Observed and model simulated seasonal cycle (number of hurricanes per month) for the N. Atlantic and E. Pacific from the 3-member ensemble mean (1=JAN, 12=DEC).

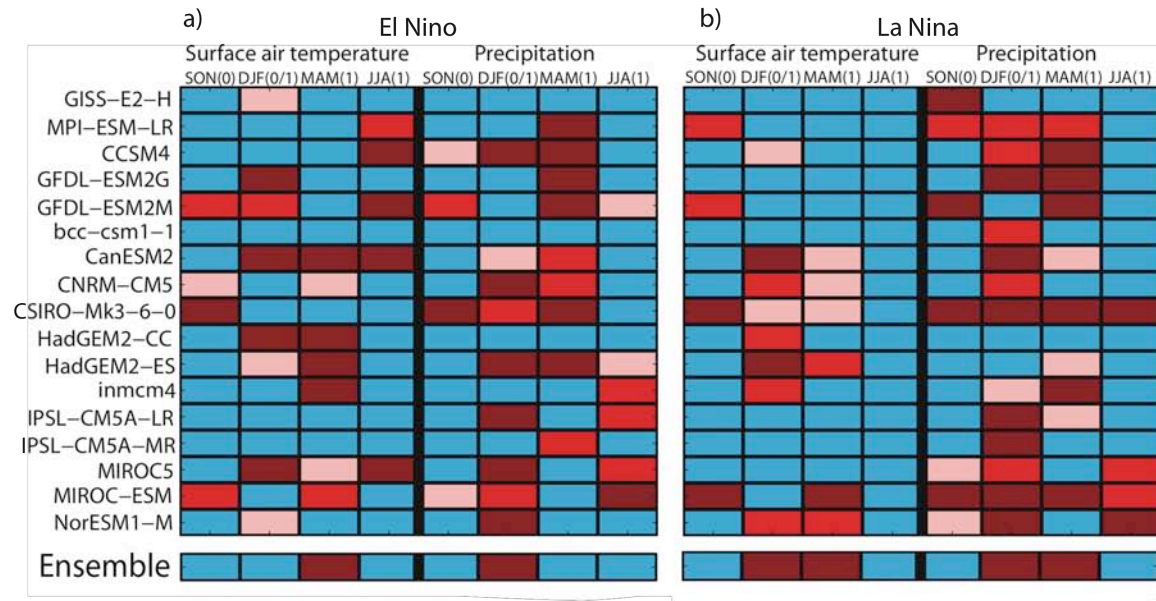
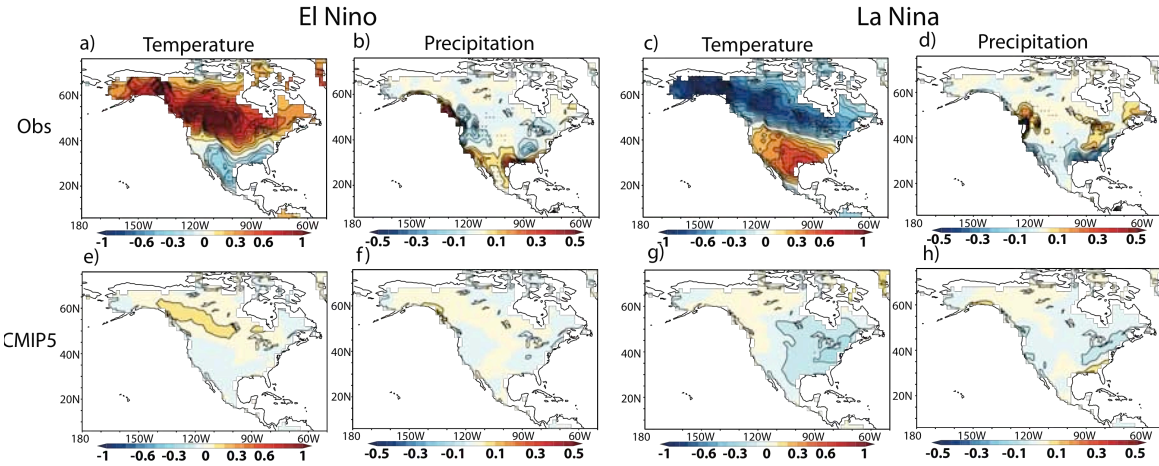


Figure 9. Comparison of ENSO-related SAT and precipitation composite patterns between CMIP5 models and observations (CRU TS3.1). Global significance of the differences between SAT (left columns) and precipitation (right columns) composites from CMIP5 models (rows) and observations for each season (designated above each column) and for (a) El Niño and (b) La Niña episodes. The model names are given at the left, and the bottom row corresponds to the CMIP5 ensemble. Light, medium, and dark red correspond with 90%, 95%, and 99% significance levels, respectively, and blue corresponds to statistically insignificant ($\alpha_{global} > 0.1$) differences. The temperature and precipitation data are regridded to a common equal-area grid with 2° latitude spacing and variable longitude spacing. We determine whether model patterns are significantly different from the observed patterns if the patterns are statistically distinguishable in a test of field significance based on the “false discovery rate” (FDR) (Benjamini and Hochberg 1995, Wilks 2006).

1593



1594

1595

1596 Figure 10. Wintertime composites of ENSO-related SAT and precipitation anomalies in
1597 observations and the CMIP5 ensemble. (a-d) Observed and (e-h) CMIP5 ensemble
1598 normalized composites of (a,c,e,g) SAT and (b,d,f,h) precipitation in DJF(0/1) for
1599 (a,b,e,f) El Niño and (c,d,g,h) La Niña episodes. The contour interval is $0.1^{\circ}\text{C }^{\circ}\text{C}^{-1}$ for the
1600 SAT composites and $0.05 \text{ mm day}^{-1} ^{\circ}\text{C}^{-1}$ for the precipitation composites, with the zero
1601 contour line omitted. In the observed composites, stippling indicates values that are
1602 statistically significant at the 95% confidence level based on a two-sided t-test.

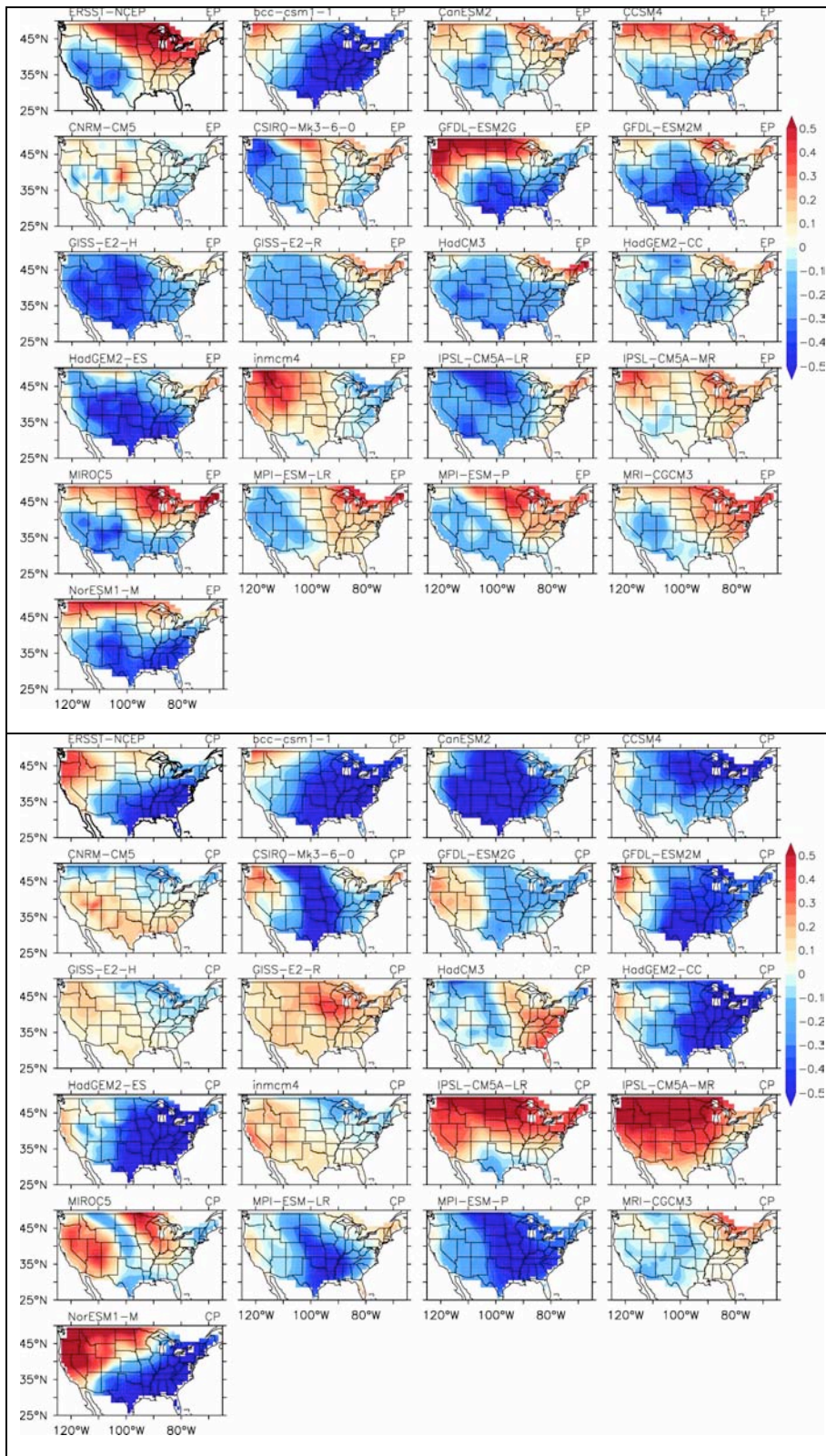


Figure 11. US winter surface air temperature regressed on the EP (top) and CP (bottom) ENSO indices from the CMIP5 models and observations (air temperature data are from the NCEP-NCAR reanalysis and SSTs are from the HadISST dataset).

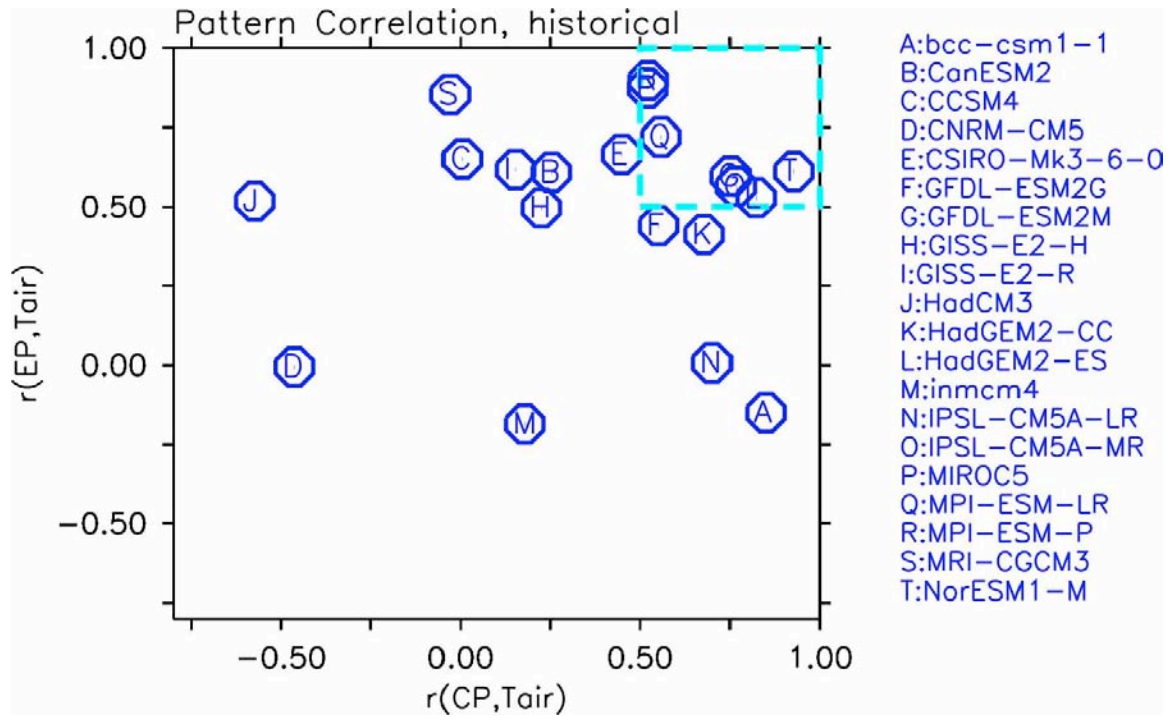


Figure 12. Scatter plot of pattern correlations between the regression patterns from the CMIP5 models and those from the observations (NCEP-NCAR reanalysis and HadISST dataset) for EP versus CP ENSO.

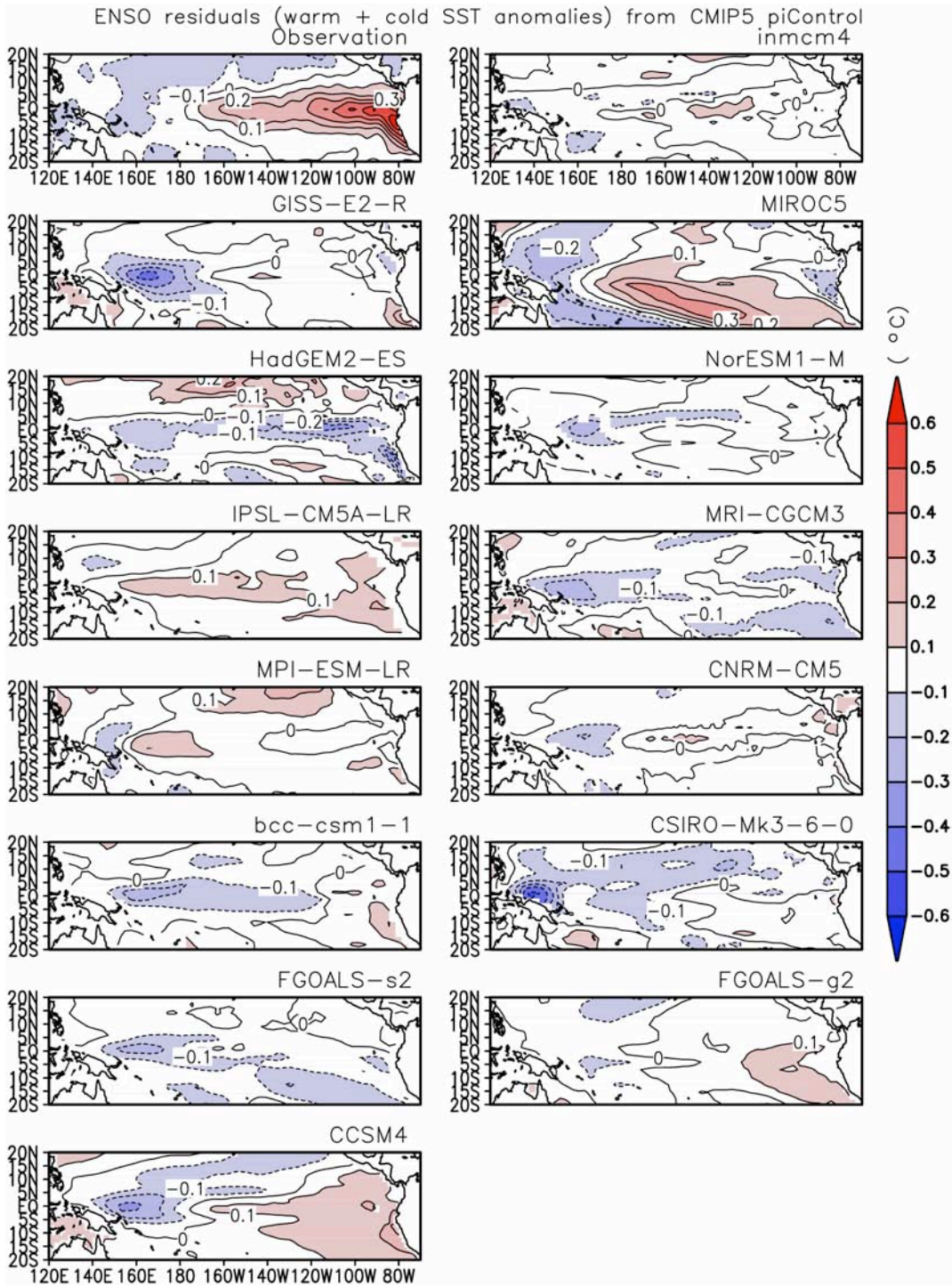


Figure 13. The sum of the composite SST anomalies between the two phases of ENSO from the HadISST observations and CMIP5 coupled models. Presented here are the results from the coupled models whose corresponding AMIP runs are available for the analysis. The definition of the warm phase and cold phase of ENSO follows that of Zhang et al. (2009). The length of data used in the calculation is 50 years for all the models and observations (1950-99).

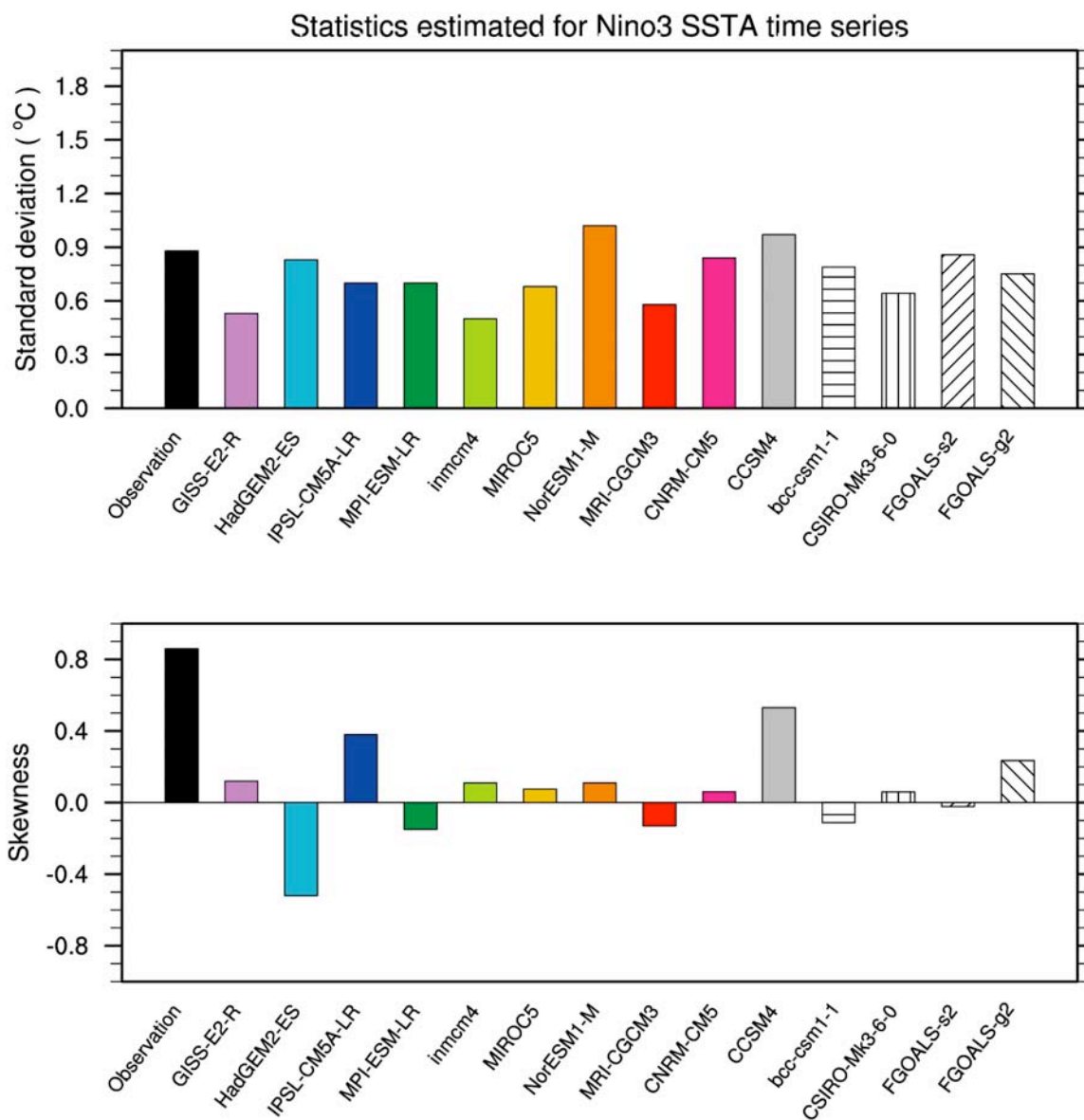


Figure 14. Standard deviation (upper) and skewness (bottom) of the interannual variability in Niño-3 SST. Data used are the same as for Figure 13.

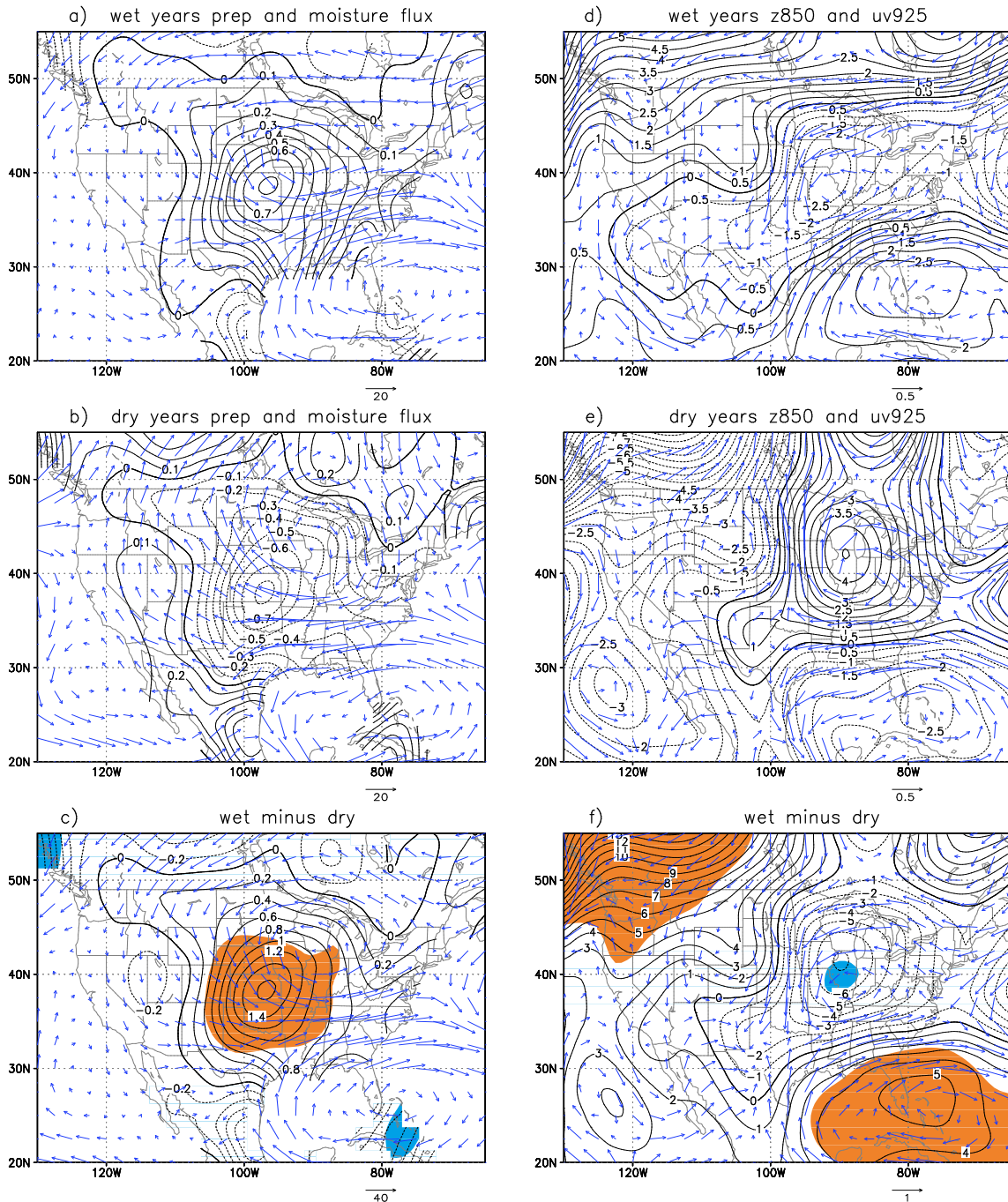


Figure 15. Reanalysis results: a) and b) show, respectively, summertime precipitation anomalies (contours) in wet and dry years, in reference to the Great Plains precipitation, and the vertically integrated moisture fluxes from the surface to the top of the troposphere (arrows). c) The differences between a) and b). d) and e) show the corresponding 850hPa geopotential height (countour) and 925hPa wind anomalies (arrows) for the wet and dry summer, respectively. Their differences are summarized in f).

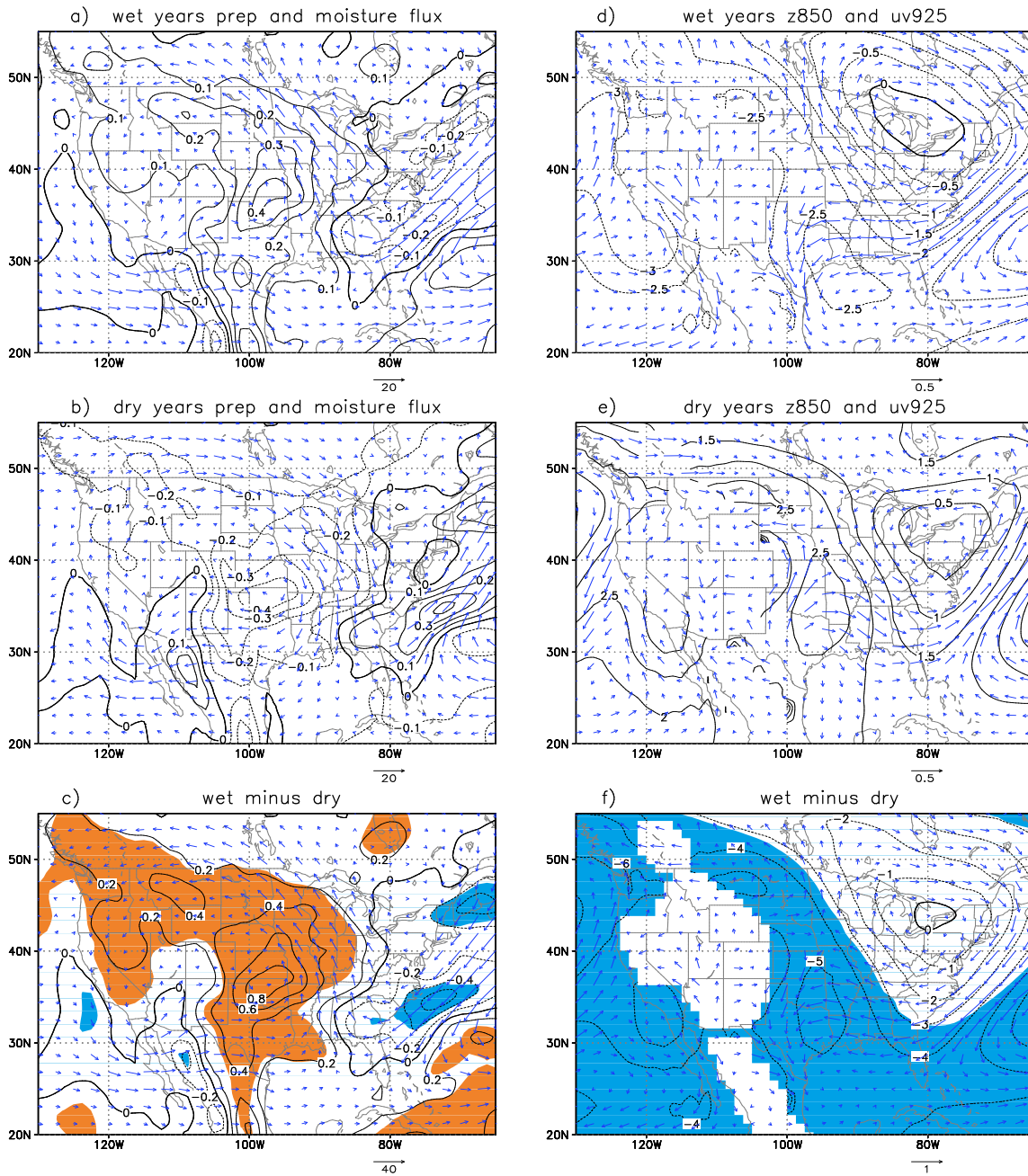


Figure 16. Same as Fig. 15 but for CCSM4 simulation results.

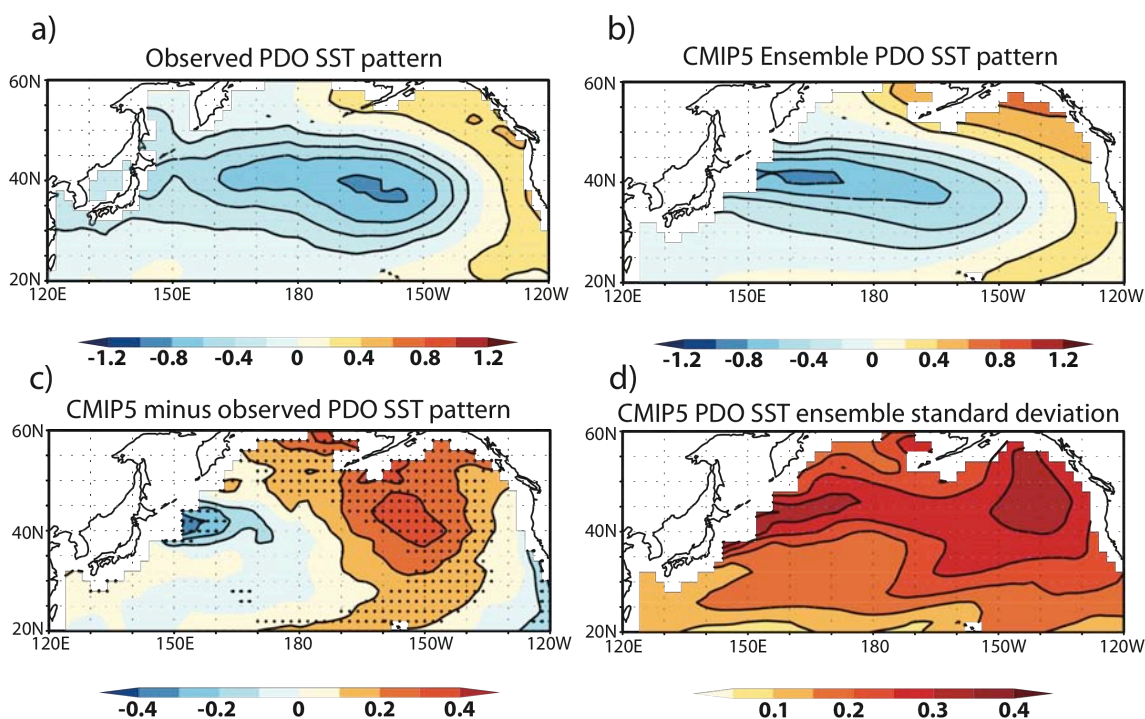


Figure 17. PDO SST patterns in observations and CMIP5 models. Linear regression of SST upon the PDO index in (a) observations and (b) the CMIP5 ensemble, and (c) the CMIP5 minus observed PDO regression. The contour interval is 0.2°C in (a) and (b) and 0.1°C in (c), with the zero contour omitted. Stippling in (c) indicates where the differences are statistically significant at the 95% confidence level based on a two-sided t-test. (d) Standard deviation of the PDO SST regressions within the ensemble. Contour interval is 0.05°C .

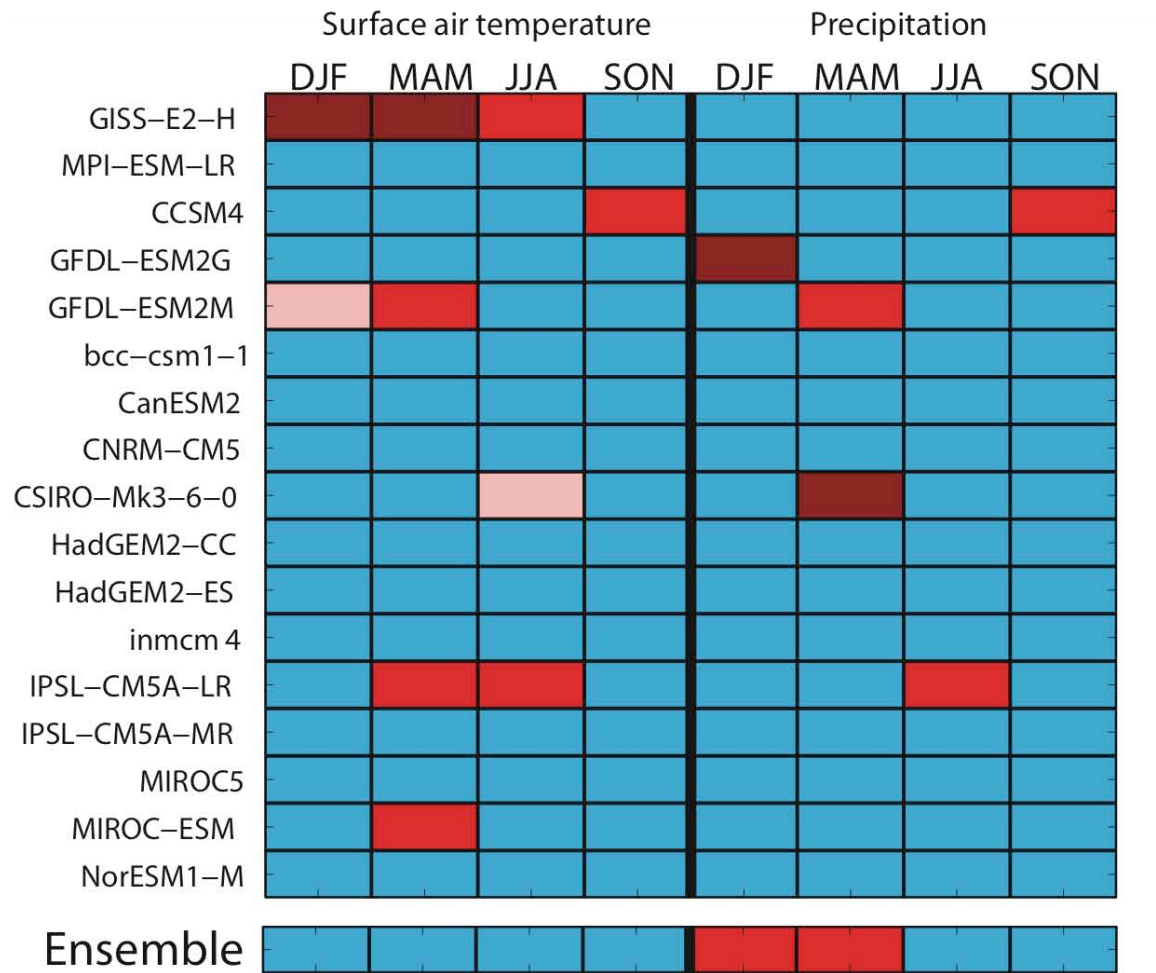


Figure 18. Comparison of PDO SAT and precipitation regression patterns between CMIP5 models and observations. Global significance of the differences between SAT (left columns) and precipitation (right columns) regressions from CMIP5 models (rows) and observations for each season. The model names are given at the left, and the bottom row corresponds to the CMIP5 ensemble. Light, medium, and dark red correspond with 90%, 95%, and 99% significance levels, respectively, and blue corresponds to statistically insignificant ($\alpha_{global} > 0.1$) differences. To highlight multi-decadal variability, a Butterworth 10-year low-pass filter is applied to each PDO index time series, which is then re-standardized. The SAT and precipitation anomalies are regressed in the North America region on this index for each season. The land temperature and precipitation datasets are from CRU TS3.1 for the period of 1901-2009 and all observation and model temperature and precipitation data are regridded to a common equal-area grid, as described in section 5.1. Prior to the regressions, the seasonal SAT and precipitation trend is removed, where the trend is obtained by a LOWESS smoothing with a span of 60 years. Significance is calculated on the regression at each grid point with a two-sided t-test, with the effective degrees of freedom adjusted for the lag-1 autocorrelation in the residuals (Santer et al., 2000).

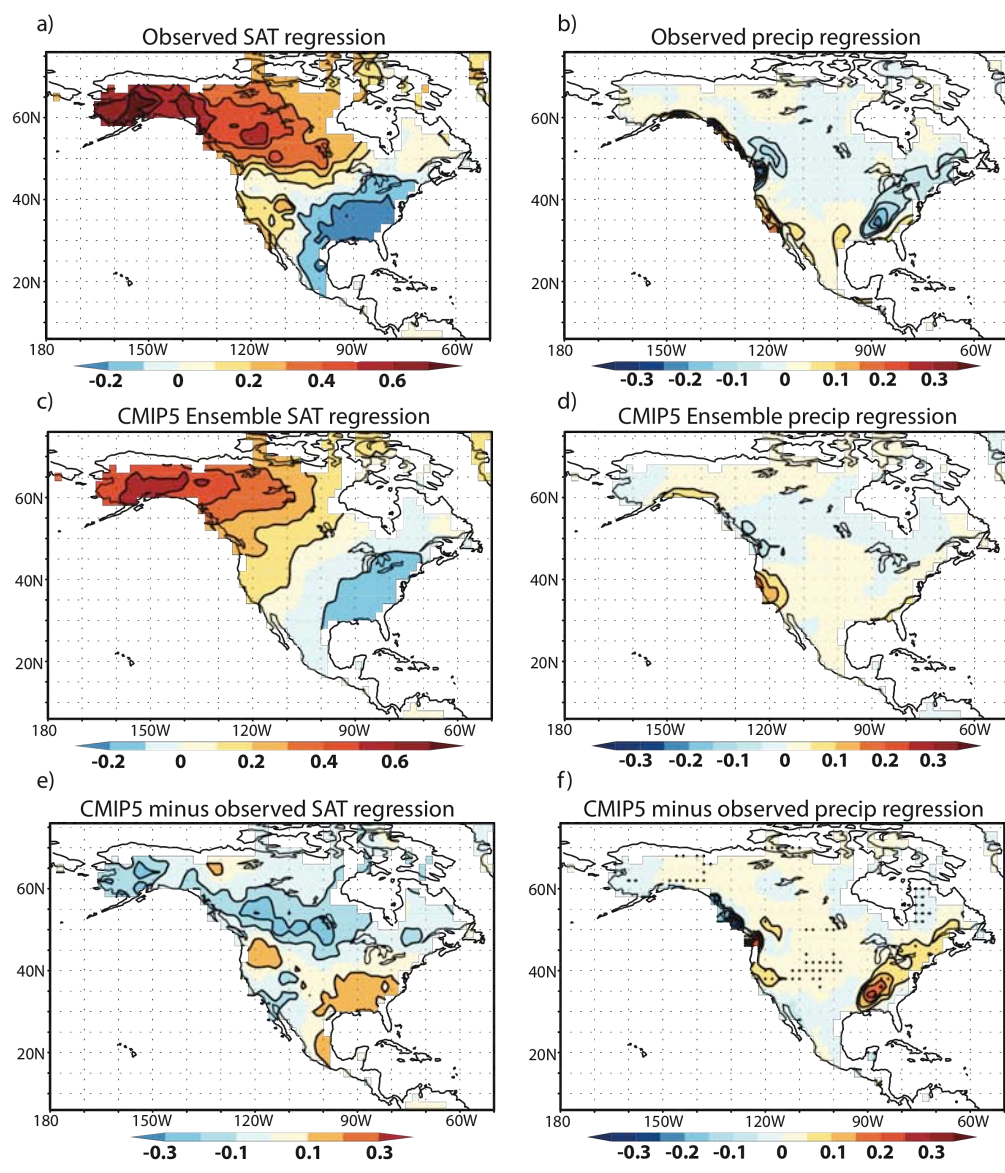


Figure 19. December-February PDO SAT and precipitation regression patterns over North America. Regressions of DJF SAT (a,c) and precipitation (b,d) upon the PDO index in (a,b) observations and (c,d) the CMIP5 ensemble. The differences between the regression patterns (CMIP5 minus observations) are shown in (e) and (f). The contour interval is 0.1°C for the SAT regressions (a, c, e) and 0.05 mm/day for the precipitation regressions (b,d,f). Stippling in (e) and (f) correspond to differences that are significantly different at the 95% confidence level based on a two-sided t-test.

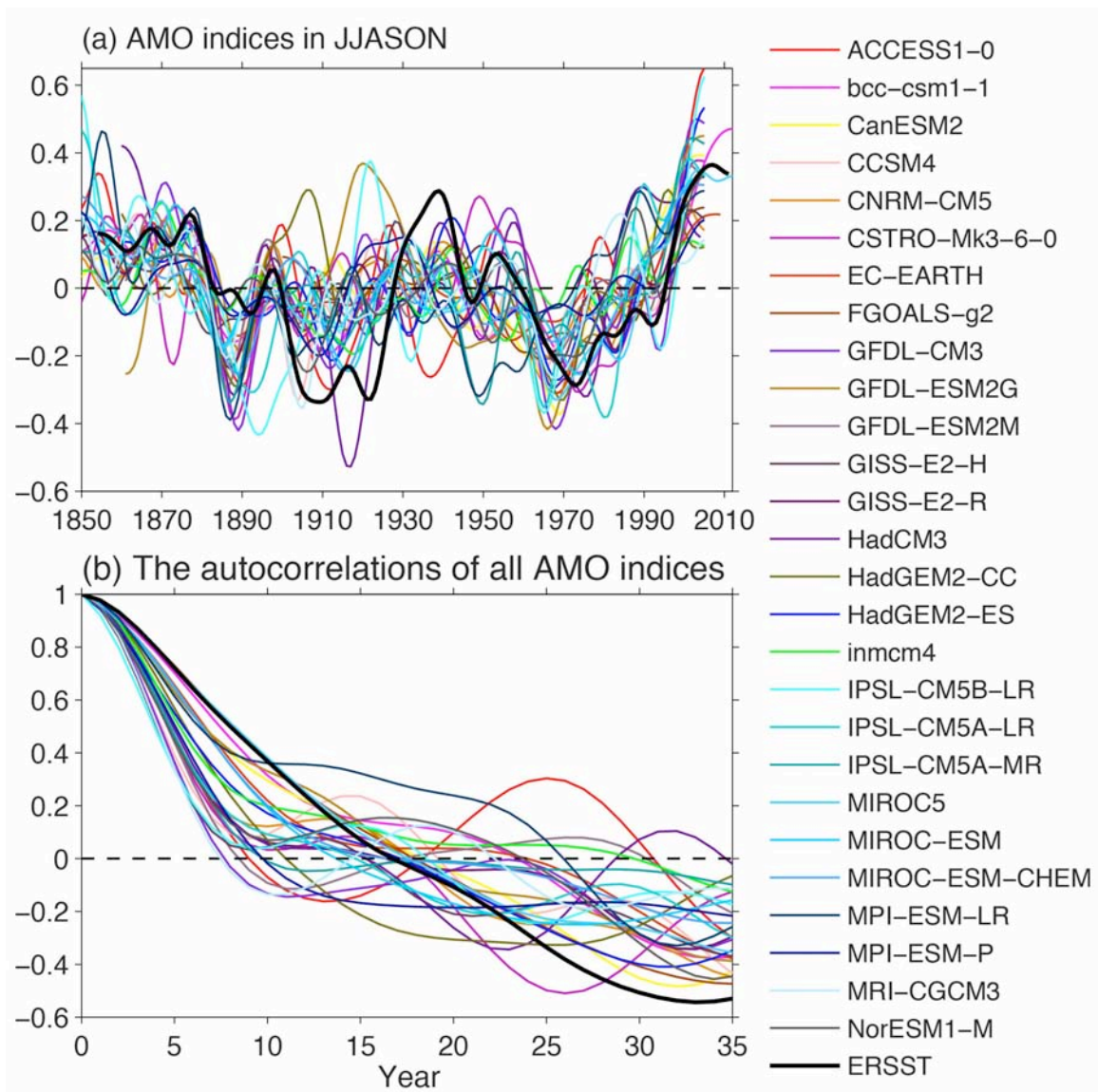


Figure 20. Time series of the AMO indices from observation and models. The AMO index defined as the detrended North Atlantic SST during the Atlantic hurricane season of June to November (JJASON) from the equator to 60°N, 75°W-5°W with the 11-year running mean. The observed AMO index is plotted in the first panel. The autocorrelations of all AMO indices are plotted in the last panel. SST is from the NOAA extended reconstructed SST version 3 (Smith et al. 2008).

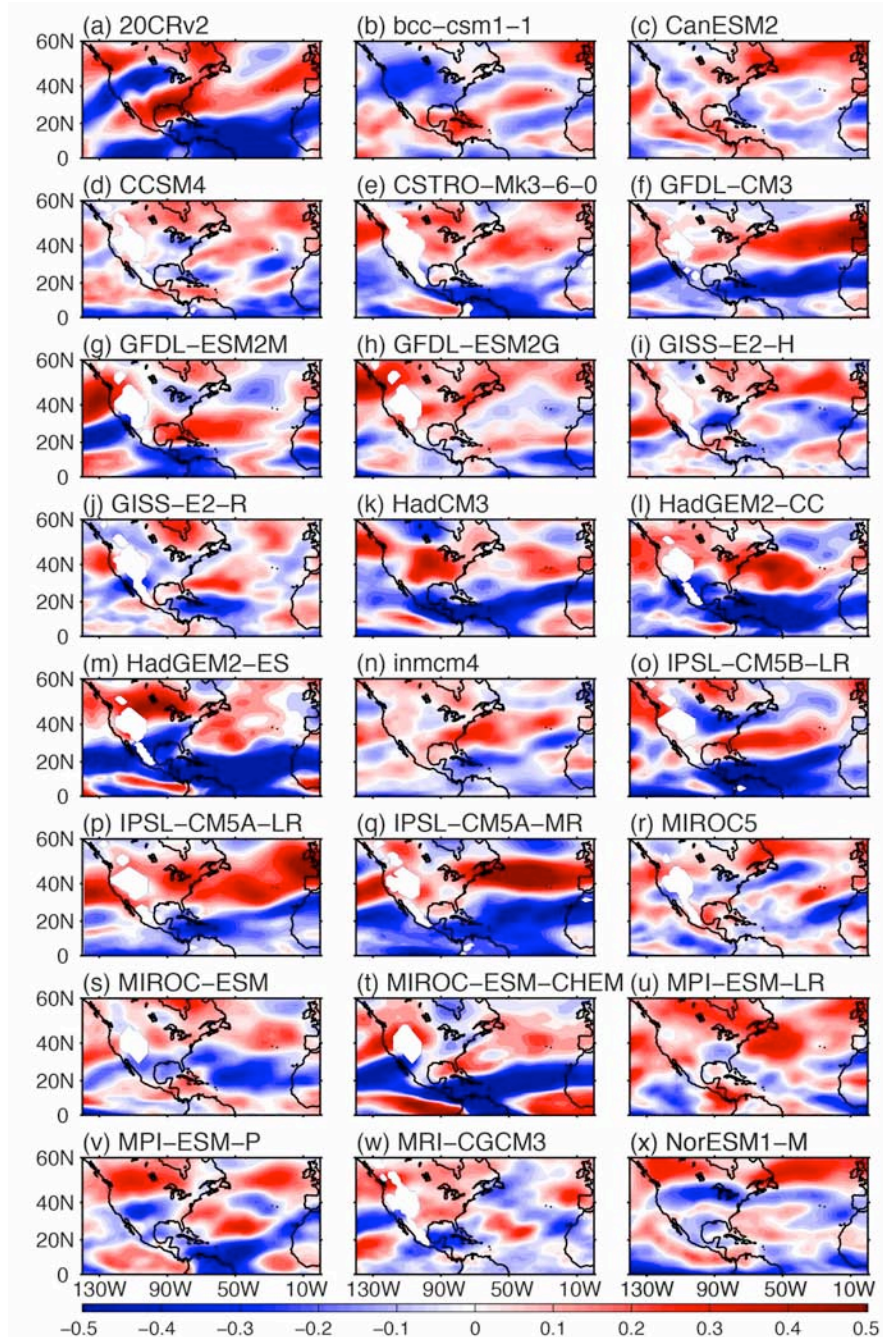


Figure 21. Regressions of the JJASON AMO on vertical wind shear. The observed regression is plotted in the first panel. The vertical wind shear is calculated as the magnitude of the vector difference between winds at 200 mb and 850 mb. The atmospheric reanalysis data set is the NOAA Earth System Research Laboratory (ESRL) 20th Century Reanalysis (20CR) (Compo et al. 2011).

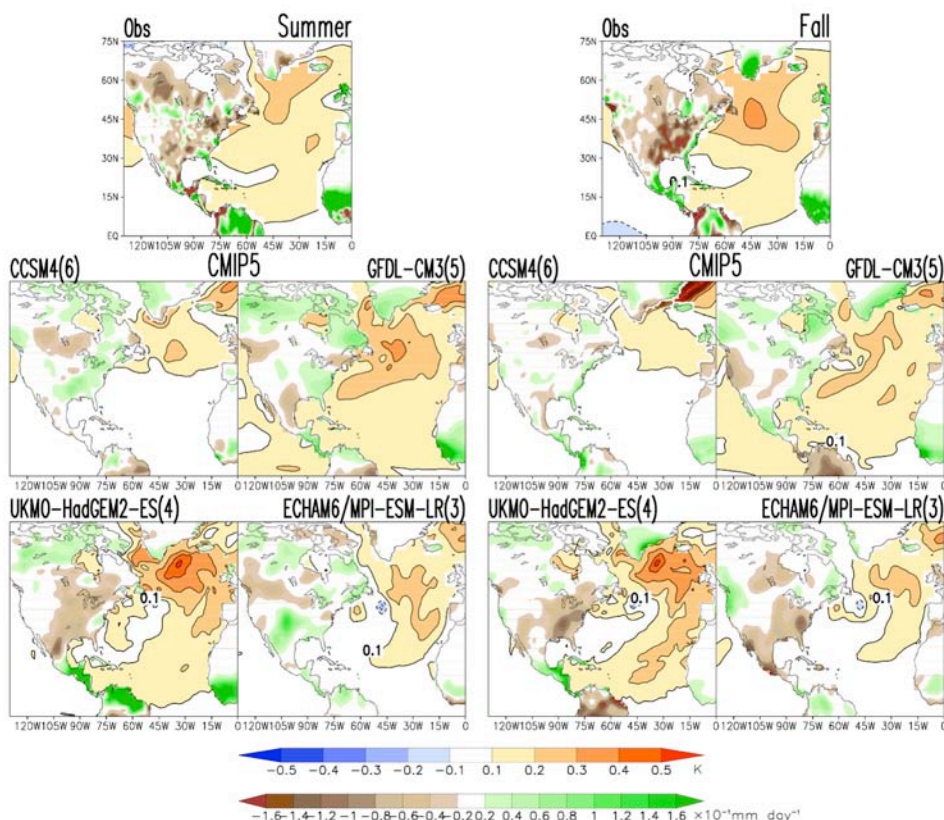
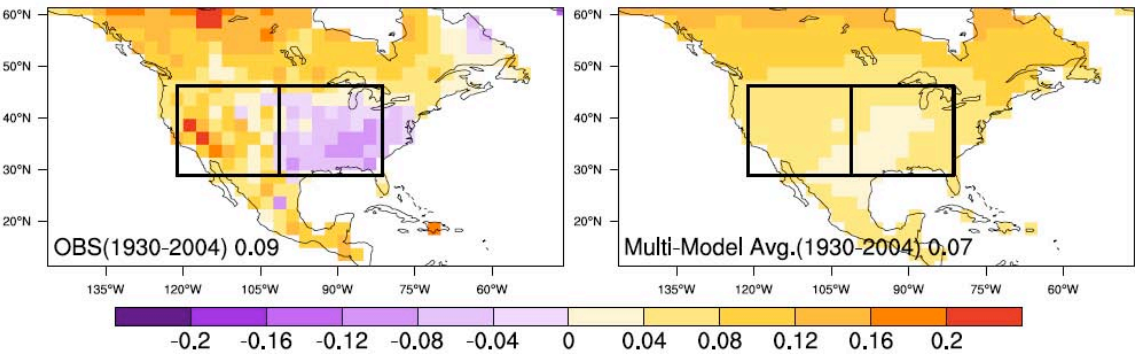


Figure 22. Summer and fall regressions of AMO indices on SST and precipitation in observations from HadISSTv1.1 and CRUTS3.1 data sets, and historical simulations of the 20th century climate from CMIP5 models for the 1901-1999 common period. AMO indices are calculated as area-averaged SST anomalies over the domain (75°-5°W, 0°-60°N) which are detrended first and then smoothed via a 11-year running mean. Regressions are calculated for each individual ensemble of a given model, then the mean of the regressions is displayed. The number in parenthesis denotes the number of ensembles used from each model. Blue/red shading denotes negative/positive SST anomalies, while brown/green shading denotes negative/positive precipitation anomalies. Contour interval is 0.1K and 0.02 mm day⁻¹.

1696



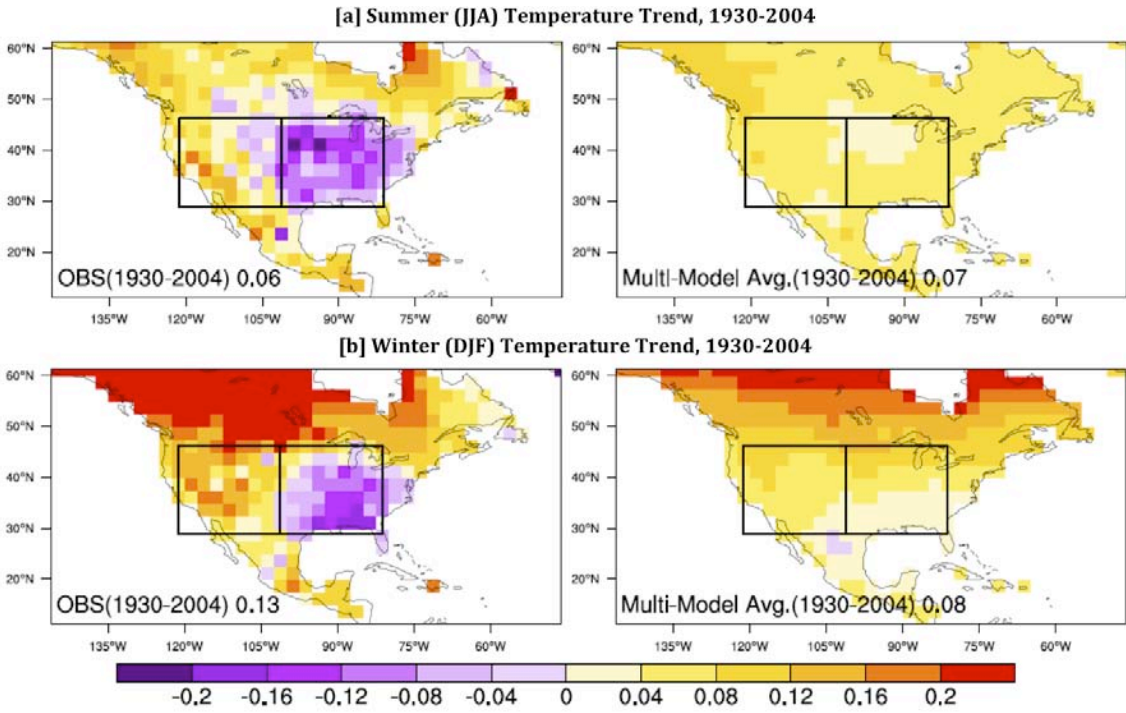
1697

1698 Figure 23. Observed and multimodel average temperature trends in North America
1699 (1930-2004). Observations are from the CRU TS3.1 dataset. The multimodel average is
1700 calculated from the best ensemble member from each model, selected based on the
1701 highest spatial correlation with observed temperature trends over North America in the
1702 respective model. Number in each panel show global mean land only temperature trend
1703 (60S to 60N). The models are CanESM2, CCM4, CNRM-CM5, CSIRO-Mk3-6-0,
1704 GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadCM3,
1705 HadGEM2-CC, HadGEM2-ES, INMCM4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-
1706 ESM, MIROC5, MPI-ESM-LR, MRI-CGCM3.

1707

1708

1708



1709

1710 Figure 24. Summer and winter temperature trend (1930-2004), unit: degree C/decade.
1711 Multi-model average is based on 19 models, best ensemble member from each model as
1712 in Fig. 23. Number in each panel show global mean land only temperature trend (60S to
1713 60N). East and West USA regions are also shown by boxes.

1714

1715

1716

1717

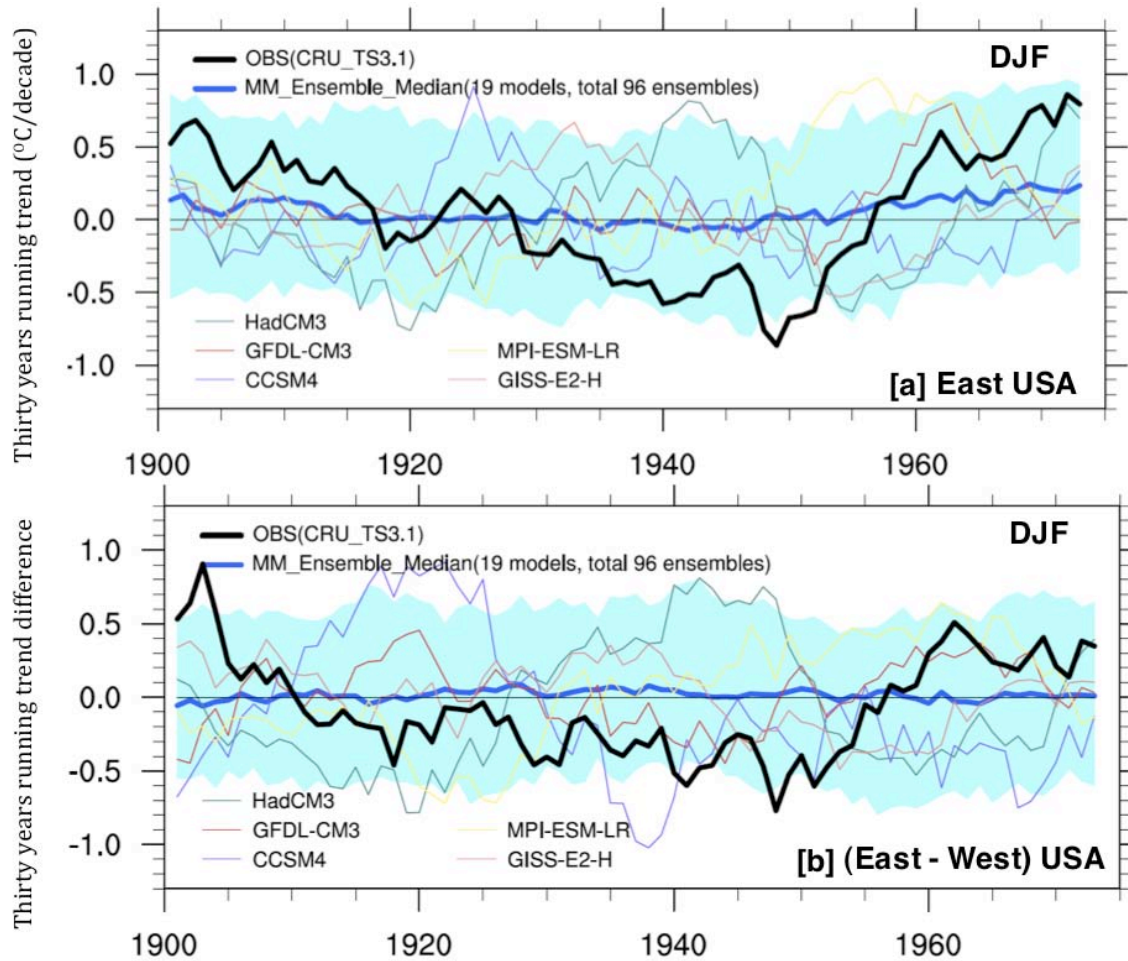


Figure 25. (a) 30 years running trend in East USA (30N to 47N, and 260E to 280E) DJF temperature trend (b) 30 year running trend difference between East and West USA. West USA is defined as (30N to 47N, and 240E to 260E). Shaded region show 95% uncertainty range calculated from total 96 ensemble members. X axis represent starting of 30 year period, for example trend corresponding to 1930 represent trend from 1930 to 1959. Five core models are also shown (only first ensemble member from each model).

Trend of Daily Temperature Range (DTR) during 1951-2000

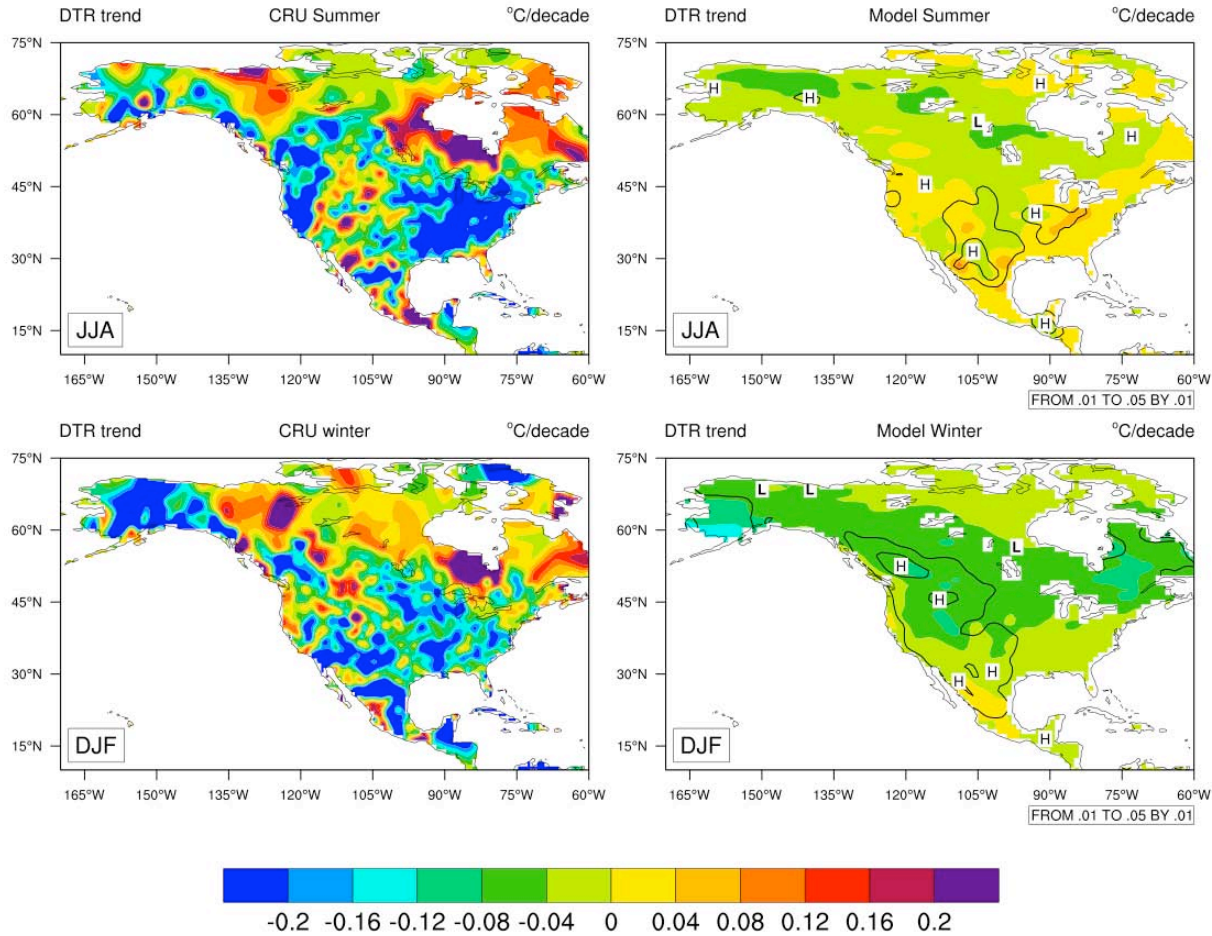
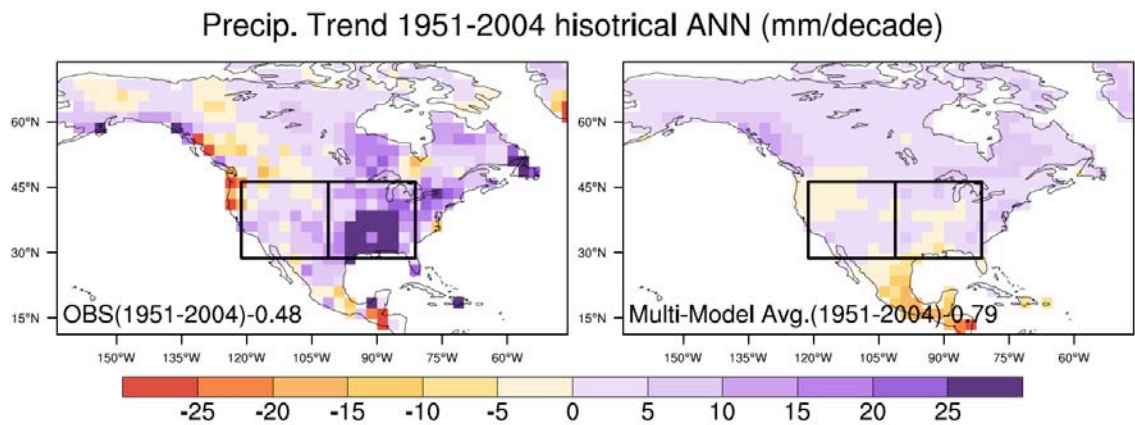


Figure 26. Comparison of observed (left) and model simulated linear trends in daily temperature range (Tmax -Tmin) during 1951-2000 period. The model ensemble mean daily temperature range (DTR) is computed from 23 models totaling 109 members available in the historical experiment. The contour lines are the inter-model variance of the trends.

1733

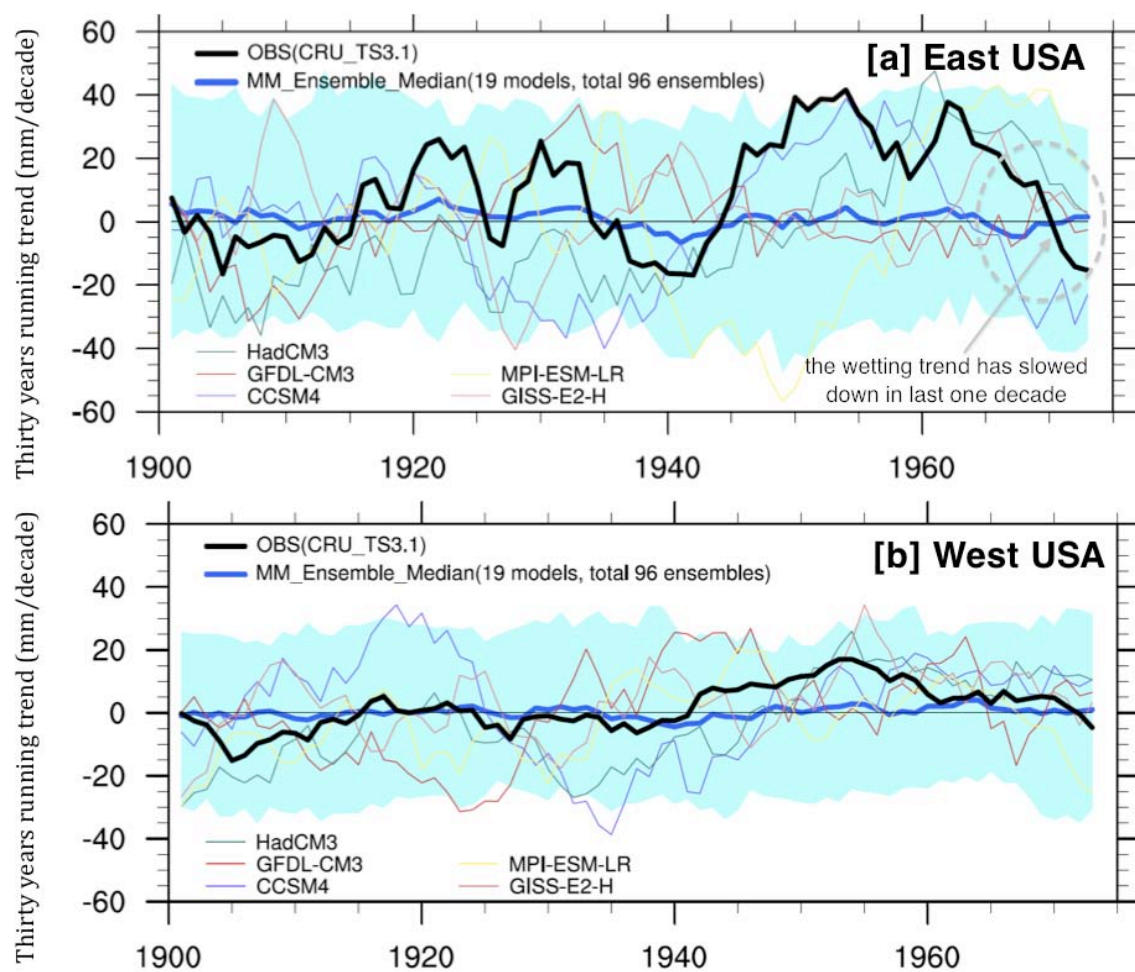


1734

1735 Figure 27. Observed and multi-model average annual precipitation trend over North
1736 America. Observation is the CRU TS3.1 dataset. The multi-model average is calculated
1737 from 19 climate models, one ensemble member from each model (see Figure 23).

1738

1738



1739

1740 Figure 28. (a) 30 years running annual precipitation trend in East USA; (b) 30 year
1741 running annual precipitation trend in West USA. East and West USA regions are shown
1742 in Fig. 27. Shaded region is the 95% uncertainty range calculated from a total of 96
1743 ensemble members. X axis represents the start year of the 30 year period. For example,
1744 the value at 1930 represents the trend from 1930 to 1959. Five example models are also
1745 shown (first ensemble member).

1746

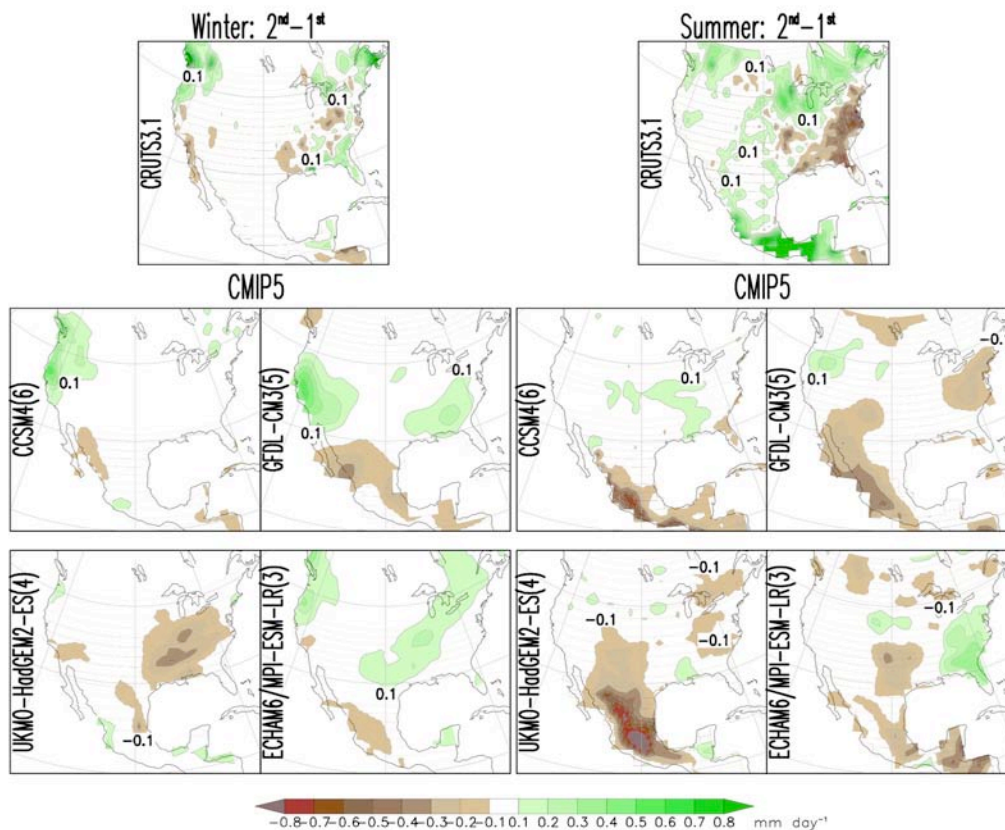


Figure 29. Differences between climatologies of the second half (1951-1999) and the first half (1901-1950) of winter and summer precipitation in observations from CRUTS3.1 data set, and historical simulations of the 20th century climate from CMIP5 models. The number in parenthesis denotes the number of ensembles used from each model. Brown/green shading denotes negative/positive precipitation anomalies; contour interval is 0.1 mm day⁻¹.